Physica D 57 (1992) 185–225
North-Holland

PHYSICA D

Orbits homoclinic to resonances, with an application to chaos in a model of the forced and damped sine-Gordon equation

Gregor Kovačič[a] and Stephen Wiggins[b, 1]

[a]*Mathematical Sciences Department, Rensselaer Polytechnic Institute, Troy, NY 12180, USA.*
[b]*Applied Mechanics 104-44, CALTECH, Pasadena CA 91125, USA.*

Received 18 December 1989
Revised manuscript received 3 September 1991
Accepted 20 February 1992
Communicated by H. Flaschka

In this paper we develop new global perturbation techniques for detecting homoclinic and heteroclinic orbits in a class of four dimensional ordinary differential equations that are perturbations of completely integrable two-degree-of-freedom Hamiltonian systems. Our methods are fundamentally different than other global perturbation methods (e.g. standard Melnikov theory) in that we are seeking orbits homoclinic and heteroclinic to fixed points that are created in a resonance resulting from the perturbation. Our methods combine the higher dimensional Melnikov theory with geometrical singular perturbation theory and the theory of foliations of invariant manifolds.

We apply our methods to a modified model of the forced and damped sine-Gordon equation developed by Bishop et al. We give explicit conditions (in terms of the system parameters) for the model to possess a symmetric pair of homoclinic orbits to a fixed point of saddle-focus type; chaotic dynamics follow from a theorem of Silnikov. This provides a mechanism for chaotic dynamics geometrically similar to that observed by Bishop et al.; namely, a random "jumping" between two spatially dependent states with an intermediate passage through a spatially independent state. However, in order for this type of Silnikov dynamics to exist we require a different, and unphysical, type of damping compared to that used by Bishop et al.

1. Introduction

The existence of a trajectory connecting an equilibrium (or *fixed*) point of an ordinary differential equation, a *homoclinic orbit*, is of special significance in virtually all applications in which such solutions arise. For example, often such trajectories lie at the heart of what is termed bursting phenomena in mathematical biology [1]. Homoclinic orbits have been shown to underlie the phenomena of intermittency in fluid mechanics [2]. There are several examples where homoclinic orbits play a role in the chaotic behavior of electrical circuits [3]. In nonlinear optics it has been shown that homoclinic orbits can be responsible for the chaotic behavior of lasers as well as light pulses in fiber optics applications [4]. Homoclinic orbits have been shown to be responsible for chaotic behavior arising in structural mechanics [5] and there are a variety of chemical reactions where chaotic oscillations in the reactant concentrations are due to the presence of a homoclinic orbit [6]. In fact, it is not an exaggeration to claim that in virtually every manifestation of chaotic behavior known thus far, some type of homoclinic behavior is lurking in the background. Thus, it is important to have mathematical methods for detecting homoclinic orbits in specific systems of ordinary differential equations.

[1]Author for correspondence.

Unfortunately, there are few analytical methods for proving the existence, or nonexistence, of homoclinic orbits in a given ordinary differential equation. For two-dimensional problems phase plane techniques can be used, however the situation becomes especially difficult for systems having phase spaces of dimension larger than two. The most widely applicable methods are the so-called *Melnikov methods* [7]. These are global perturbation methods that utilize the geometrical features of integrable Hamiltonian systems as a framework on which to develop perturbation methods. Roughly speaking, the unperturbed system is an integrable Hamiltonian system having a normally hyperbolic invariant set whose stable and unstable manifolds intersect nontransversely. The global geometry associated with the integrable structure is used to develop coordinates which are used in determining if any of the homoclinic orbits to the normally hyperbolic invariant set (which is affected only slightly by the perturbation) survive under perturbation.

In this paper we develop a new method for proving the existence of homoclinic orbits in a class of perturbed, integrable two-degree-of-freedom Hamiltonian systems, where the perturbation is not Hamiltonian. Our methods are fundamentally different from previous methods in that we prove the existence of orbits homoclinic to fixed points that are created by the perturbation in a resonance. More precisely, in the four dimensional phase space there is a two-dimensional normally hyperbolic locally invariant manifold which, for the unperturbed problem, contains a circle of fixed points. This circle of fixed points arises as a result of the vanishing of a frequency, which is the origin of the term *resonance*. Under the perturbation this circle of fixed points "blows up" into the typical resonance band structure, restricted to the two dimensional locally invariant manifold. Namely, a pair of fixed points survive, and become a saddle and a sink, with the unstable manifold of the saddle falling into the sink and the stable manifold of the saddle forming the boundary of the basin of attraction of the sink. We then develop a method for determining the existence of orbits homoclinic to the fixed point in the resonance that is a sink when restricted to the two-dimensional manifold.

Our method involves two steps: in the first step we show the existence of an orbit that leaves the fixed point and returns to a neighborhood of the resonance using an argument very much like that used in the standard higher dimensional Melnikov theory. In the next step we use the theory of fibering of stable and unstable manifolds by submanifolds corresponding to initial conditions of trajectories that have the same asymptotic phase in order to show that the orbit actually returns to the fixed point. This last step is essentially a problem in singular perturbation theory and these *stable fibers* enable us to transform it to a problem in regular perturbation theory.

Our methods apply to systems of the following form:

$$\dot{x} = J\mathrm{D}_x H(x,I) + \epsilon g^x(x,I,\gamma,\mu,\epsilon), \quad \dot{I} = \epsilon g^I(x,I,\gamma,\mu,\epsilon), \quad \dot{\gamma} = \mathrm{D}_I H(x,I) + \epsilon g^\gamma(x,I,\gamma,\mu,\epsilon)$$
$$(x,I,\gamma) \in \mathbb{R}^2 \times \mathbb{R} \times \mathrm{S}^1, \quad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad (1.1)_\epsilon$$

where $\mu \in \mathbb{R}^p$ is a vector of parameters and $0 < \epsilon \ll 1$ (*note:* throughout this paper "D_x", etc. denotes the partial derivative with respect to x, etc. whereas $\mathrm{d}/\mathrm{d}x$, etc. will denote the total derivative with respect to x, etc.). In section 2 we describe our assumptions on the structure of the unperturbed system (i.e., $(1.1)_\epsilon$ with $\epsilon = 0$), in particular, we describe the degenerate structure that leads to orbits homoclinic to a resonance band. In section 3 we describe certain results from the theory of normally hyperbolic invariant manifolds that give us some general information about the structure of the perturbed equations, especially the regular versus singular perturbation aspects and the geometrical interpretation. In section 4 we develop the method for determining the existence of homoclinic connections and in

section 5 we discuss conditions under which these homoclinic connections can lead to chaotic dynamics. In section 6 we apply the theory to a set of ordinary differential equations that model certain modal interactions in the perturbed nonlinear Schrödinger equation. For completeness we give some background for this particular application.

The development of all of the theory in this paper was motivated by the desire to understand a variety of numerical experiments on the damped and driven sine-Gordon equation performed by Bishop, Ercolani, Forest, McLaughlin, and Overman [8]. We give a brief description of their results. Consider the perturbed sine-Gordon equation

$$u_{tt} - u_{xx} + \sin u = \epsilon\left(-\hat{\alpha}u_t + \hat{\Lambda}u_{txx} + \hat{\Gamma}\sin \omega t\right),$$

with periodic, even boundary conditions:

$$u\left(x = -\tfrac{1}{2}L, t\right) = u\left(x = \tfrac{1}{2}L, t\right), \quad u(x,t) = u(-x,t),$$

with $0 < \epsilon\hat{\alpha} \ll 1$, $0 < \epsilon|\hat{\Lambda}| \ll 1$, $\omega = 1 - \epsilon\tilde{\omega}$, and L fixed where ϵ and $\tilde{\omega}$ are both positive. In the original numerical experiments $\hat{\Lambda}$ was taken to be zero, however we will include it in our analysis and discuss the reasons for this when we describe the *two-mode model* at the end of the introduction. In the numerical experiments Bishop et al. chose a single hump sine-Gordon "breather" as an initial condition and observed its evolution in space and time as they increased the forcing $\epsilon\hat{\Gamma}$.

The complete bifurcation sequence is described in ref. [9]. Here we only present the main point, which is that above a certain value of $\epsilon\hat{\Gamma}$ (with $\hat{\Lambda} = 0$) chaotic jumping of the solution occurs between two "breathers", one peaked at the middle and the other at the ends of the interval $[-\frac{1}{2}L, \frac{1}{2}L]$, with the solution passing near a spatially independent (or "flat") solution on every jump. Comparing this situation with the geometrical structure in the phase space of the *unperturbed* sine-Gordon equation [10], they found that the latter possesses linearly unstable spatially independent solutions connected to themselves by homoclinic orbits: the two types of "breathers" with a spatial hump structure which exhibit chaotic behavior in the perturbed problem. From the knowledge of the geometrical structure of the phase space of the unperturbed system, as well as the fact that homoclinic orbits appeared to be involved in the numerically observed chaos, Bishop et al. inferred that a Melnikov-type analysis could possibly be developed in order to see how these unperturbed structures are distorted under perturbation in such a way as to give rise to chaos. As a preliminary step in the analysis they chose to develop a simple model that captures the essential structure of the unperturbed and perturbed sine-Gordon equation. It is this simple model that we now describe.

At small amplitudes, and for frequencies close to 1, the dynamics of the sine-Gordon equation can be approximated by the nonlinear Schrödinger envelope equation [11]. This can be seen as follows; if we seek a solution of the perturbed sine-Gordon equation of the form

$$u_\epsilon(x,t) = 2\sqrt{\epsilon\tilde{\omega}}\left[B(X,T)\,e^{i\omega t} + B^*(X,T)\,e^{-i\omega t}\right] + \mathcal{O}(\epsilon),$$

with $X = \sqrt{2\epsilon\tilde{\omega}}\,x$ and $T = \epsilon\tilde{\omega}t$, and substitute it into the perturbed sine-Gordon equation we obtain the following perturbed nonlinear Schrödinger equation for $B(X,T)$:

$$-iB_T + B_{XX} + (|B|^2 - 1)B = \epsilon\left(i\alpha B - i\Lambda B_{XX} + i\bar{\Gamma}\right),$$

where $\hat{\alpha} = 2\epsilon\tilde{\omega}\alpha$, $\hat{\Lambda} = \Lambda$, and $\hat{\Gamma} = 8\epsilon^{3/2}\hat{\omega}^{3/2}\bar{\Gamma}$. This equation is defined on the X interval $[-\frac{1}{2}L_X, \frac{1}{2}L_X]$ where $L_X = \sqrt{2\epsilon\tilde{\omega}}\, L$.

Bishop et al. next constructed a two (Fourier) mode truncation to the nonlinear Schrödinger equation by assuming a solution of the form

$$B(X,T) = \frac{1}{\sqrt{2}}c(T) + b(T)\cos kX$$

(with $k = 2\pi/L_X$). Substituting this solution into the equation and neglecting the higher fourier modes gives the following *two-mode model*

$$-\mathrm{i}\dot{c} + \left(\tfrac{1}{2}|c|^2 + \tfrac{1}{2}|b|^2 - 1\right)c + \tfrac{1}{2}(cb^* + bc^*)b = \mathrm{i}\epsilon\alpha c + \mathrm{i}\epsilon\Gamma,$$

$$-\mathrm{i}\dot{b} + \left[\tfrac{1}{2}|c|^2 + \tfrac{3}{4}|b|^2 - (1+k^2)\right]b + \tfrac{1}{2}(cb^* + bc^*)c = \mathrm{i}\epsilon\beta b,$$

where we have set $\Gamma = \sqrt{2}\,\bar{\Gamma}$, and $\beta = (\alpha + \Lambda k^2)$. Now we return to the issue of damping in the nonlinear Schrödinger equation. If we think of the solution as being decomposed into fourier modes, the term $\mathrm{i}\alpha B$ damps each mode at the same rate, however the term $-\mathrm{i}\Lambda B_{XX}$ damps modes at different rates (one can easily see the manifestation of this in the two mode model). In the two-mode model we found that the type of homoclinic orbit studied in this paper *will not exist* if the two modes are damped at equal rates. Moreover, we found that in order for the homoclinic orbit to exist the b mode must be *negatively* damped. In section 6 we present a study of the dynamics of the two-mode model where we will discuss these issues more fully.

2. Structure of the unperturbed system

As mentioned in section 1, the system obtained by setting $\epsilon = 0$ in $(1.1)_\epsilon$ is referred to as the unperturbed system which we write below:

$$\dot{x} = J\mathrm{D}_x H(x,I), \quad \dot{I} = 0, \quad \dot{\gamma} = \mathrm{D}_I H(x,I),$$
$$(x, I, \gamma) \in \mathbb{R}^2 \times \mathbb{R} \times \mathrm{S}^1. \qquad (2.1)_0$$

Note the structure of $(2.1)_0$; the x component of $(2.1)_0$ decouples from the I and γ components of $(2.1)_0$. The dynamics of the I and γ components of $(2.1)_0$ are quite simple and we make the following assumption on the dynamics of the x component of $(2.1)_0$.

Assumption 1. For all $I \in [I_1, I_2]$, the equation

$$\dot{x} = J\mathrm{D}_x H(x,I) \qquad (2.1)_{0,x}$$

has a hyperbolic fixed point, $\tilde{x}_0(I)$, connected to itself by a homoclinic orbit, $x^h(t,I)$, i.e., $\lim_{t\to\pm\infty} x^h(t,I) = \tilde{x}_0(I)$.

From $(2.1)_{0,x}$ and assumption 1 we can draw the following conclusions concerning the phase space structure of $(2.1)_0$ in the full four dimensional phase space.

2.1. The normally hyperbolic invariant manifold $\mathcal{M}$

It is easy to verify that the set of points given by

$$\mathcal{M} = \{(x, I, \gamma) | x = \tilde{x}_0(I),\ I_1 \leq I \leq I_2\} \tag{2.2}$$

is a two-dimensional, normally hyperbolic invariant manifold with boundary. The boundary of $\mathcal{M}$, denoted $\partial\mathcal{M}$, is given by the two circles $\{(x, I, \gamma) | x = \tilde{x}_0(I),\ I = I_1\}$, $\{(x, I, \gamma) | x = \tilde{x}_0(I),\ I = I_2\}$. Thus, since $\dot{I} = 0$, the vector field $(2.1)_0$ is identically zero on $\partial\mathcal{M}$. Invariance follows from assumption 1 which says that the vector field $(2.1)_0$ is tangent to $\mathcal{M}$. The term *normally hyperbolic* means that the rate of expansion and contraction under the linearized dynamics in directions complementary to $\mathcal{M}$ dominates that in directions tangent to $\mathcal{M}$. An analytical formulation as well as a proof of this statement for systems of the form of $(2.1)_0$ can be found in ref. [7]. The important point for us is that normally hyperbolic invariant sets persist under perturbation.

2.2. The dynamics on $\mathcal{M}$

The unperturbed system restricted to $\mathcal{M}$ is given by

$$\dot{I} = 0, \quad \dot{\gamma} = \mathrm{D}_I H(\tilde{x}_0(I), I), \qquad I_1 \leq I \leq I_2. \tag{2.3}$$

This system can be trivially solved and the trajectories are given by

$$I = \text{constant}, \quad \gamma(t) = \big(\mathrm{D}_I H(\tilde{x}_0(I), I)\big)t + \gamma_0. \tag{2.4}$$

Thus if $\mathrm{D}_I H(\tilde{x}_0(I), I) \neq 0$ then $I = \text{constant}$ labels a periodic orbit and if $\mathrm{D}_I H(\tilde{x}_0(I), I) = 0$ then $I = \text{constant}$ labels a circle of fixed points. We refer to a value of I for which $\mathrm{D}_I H(\tilde{x}_0(I), I) = 0$ as a *resonant I value* and we make the following assumption on the unperturbed system restricted to $\mathcal{M}$.

Assumption 2. (*Resonance*) There exists a value of $I \in [I_1, I_2]$, denoted I^r, at which $\mathrm{D}_I H(\tilde{x}_0(I^r), I^r) = 0$.

We remark that if there exists more than one resonant I value then our theory can be applied to each one individually.

2.3. $W^s(\mathcal{M})$, $W^u(\mathcal{M})$, and the homoclinic manifold, Γ

The two-dimensional normally hyperbolic invariant manifold with boundary $\mathcal{M}$ has three dimensional stable and unstable manifolds which we denote as $W^s(\mathcal{M})$ and $W^u(\mathcal{M})$, respectively. This can be inferred from the structure of the x component of $(2.1)_0$ given in assumption 1 as well as the proof of normal hyperbolicity of $\mathcal{M}$ given in ref. [7]. Moreover, the existence of the homoclinic orbit of $(2.1)_{0,x}$ implies that $W^s(\mathcal{M})$ and $W^u(\mathcal{M})$ intersect (nontransversely) along a three dimensional homoclinic

manifold which we denote by Γ. A trajectory in $\Gamma \equiv W^s(\mathcal{M}) \cap W^u(\mathcal{M})$ can be expressed as

$$\left(x^h(t,I), I, \gamma(t,I,\gamma_0) = \int_0^t D_I H(x^h(s,I),I)\,ds + \gamma_0\right) \tag{2.5}$$

and it is clear that this trajectory approaches an orbit in $\mathcal{M}$ as $t \to \pm\infty$ since $x^h(t,I) \to \tilde{x}_0(I)$ as $t \to \pm\infty$. It will be useful later on to have a parametrization of Γ so that we can describe specific points in Γ. This can be accomplished by using the fact that since $(2.1)_0$ is an autonomous system shifting the t argument in the expression for a trajectory of $(2.1)_0$ results in a new trajectory of the system. Hence

$$\left(x^h(t-t_0,I), I, \gamma(t-t_0,I,\gamma_0)\right) \tag{2.6}$$

is also a trajectory of $(2.1)_0$. Now consider the x component of $(2.1)_0$ and let $x^h(0,I)$ be a reference point on the homoclinic trajectory to $\tilde{x}_0(I)$. Then, by uniqueness of solutions, t_0 is the unique time of flight for the point $x^h(-t_0,I)$ to reach the reference point $x^h(0,I)$. Hence, in this manner, all points on the homoclinic orbit $x^h(t,I)$ to $\tilde{x}_0(I)$ can be put into 1–1 correspondence with the points in $\mathbb{R}$. Using this information for the full four-dimensional system, the homoclinic manifold Γ can be parametrized as follows:

$$\Gamma = \left\{(x,I,\gamma) \mid x = x^h(-t_0,I),\ \gamma = \gamma(-t_0,I,\gamma_0),\ I_1 \le I \le I_2,\ t_0 \in \mathbb{R},\ 0 \le \gamma_0 \le 2\pi\right\}, \tag{2.7}$$

i.e., the values of t_0, I, and γ_0 given in (2.7) can be put in 1–1 correspondence with the points in Γ via the rule specified in (2.7). We note that this parametrization of the homoclinic manifold implies that varying t_0 moves one along the homoclinic manifold *along the same orbit*. This is different from the parametrization used in [7], a comparison can be found in [15].

2.4. *The dynamics on Γ and its relation to the dynamics in $\mathcal{M}$*

Recall the expression for an orbit in Γ given in (2.5). As $x^h(t,I) \to \tilde{x}_0(I)$ and I remains constant, we want to call attention to the expression that we will define as

$$\Delta\gamma = \gamma(+\infty, I, \gamma_0) - \gamma(-\infty, I, \gamma_0). \tag{2.8}$$

Now for I such that $D_I H(\tilde{x}_0(I), I) \neq 0$ it is easy to see that $\Delta\gamma$ is not defined. This just reflects the fact that asymptotically the orbit approaches a periodic orbit whose phase constantly changes forever. However at resonant I values $\Delta\gamma$ is defined since the integral converges (convergence of the integral follows from the fact that $x^h(t,I) \to \tilde{x}_0(I)$ exponentially fast as $t \to \pm\infty$, hence at resonance $D_I H(x^h(t,I),I)$ goes to zero exponentially fast as t goes to $\pm\infty$). Since resonant I values denote circles of fixed points on $\mathcal{M}$ the orbit $(x^h(t,I^r), I^r, \gamma(t,I^r,\gamma_0))$ is typically a heteroclinic connection between different points on the resonant circle of fixed points (the connection will be homoclinic if $\Delta\gamma = 2\pi n$, for some integer n). The number $\Delta\gamma$ gives the shift in phase between the two endpoints of the heteroclinic trajectory along the circle of fixed points. In fig. 2.1 we illustrate the dynamics in the full four-dimensional phase space for the unperturbed system $(2.1)_0$.

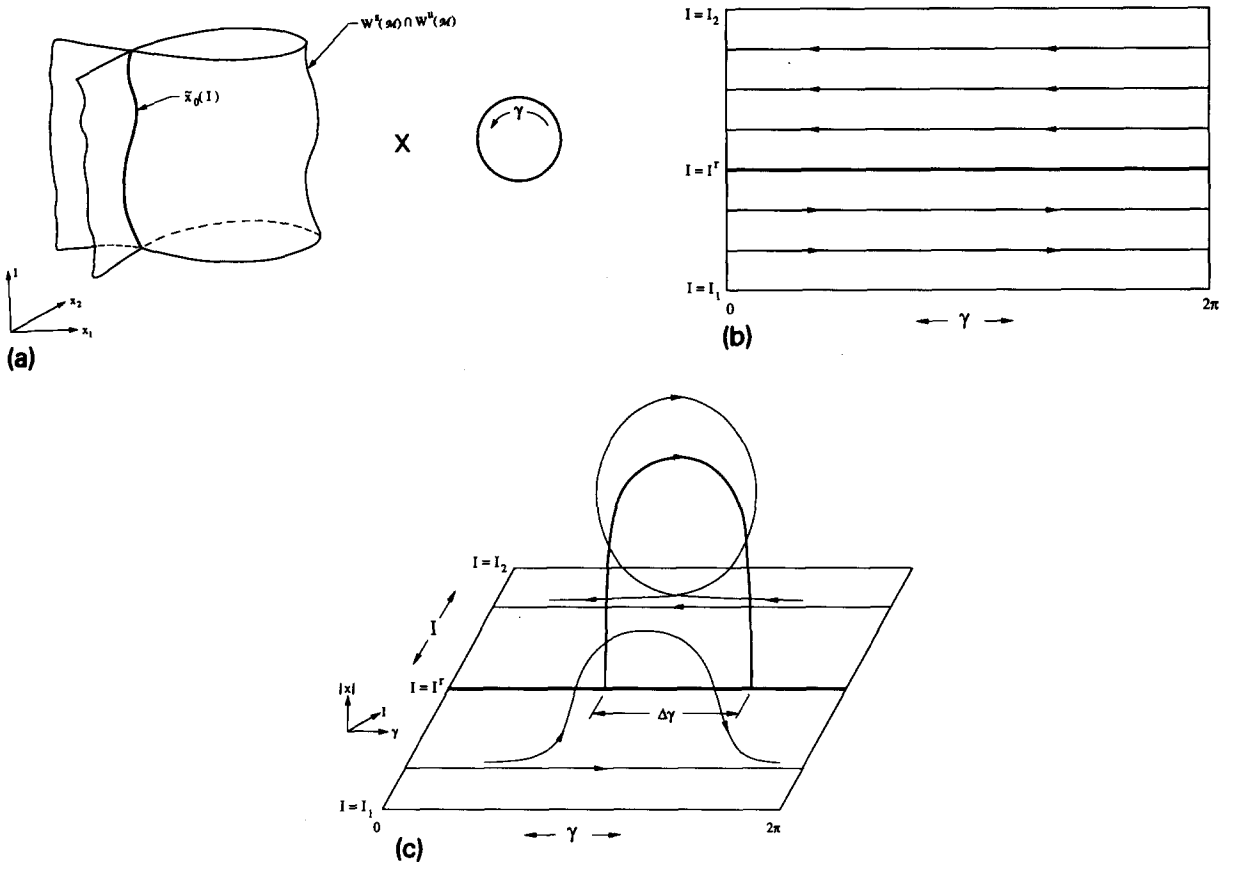

Fig. 2.1. (a) Geometry of $\mathcal{M}$ and the stable and unstable manifolds of $\mathcal{M}$. (b) The dynamics of $\mathcal{M}$. (c) The geometry of trajectories homoclinic to the periodic orbits on $\mathcal{M}$ and orbits heteroclinic to fixed points on the resonance.

3. Structure of the perturbed system

The invariant manifolds and their associated structure in the integrable unperturbed system $(2.1)_0$ described in section 2 will form the framework on which we develop our analytical methods. In particular, $\mathcal{M}$ and its stable and unstable manifolds persist under sufficiently small perturbations; however the dynamics within these manifolds will generally undergo radical changes under perturbation. In this section we will discuss these issues. As a preliminary warning to the reader we remark that we will use two different sets of coordinates in our analysis; the (x, I, γ) coordinates already defined and (x, h, γ) coordinates which will be defined shortly. We will comment for the need and reasoning for this as we go along. We begin by considering the behavior of $\mathcal{M}$ along with its stable and unstable manifolds under perturbation.

3.1. *The persistence of $\mathcal{M}$, $W^s(\mathcal{M})$, and $W^u(\mathcal{M})$ under perturbation*

The question of the persistence under perturbations of invariant manifolds with boundary gives rise to certain technical questions concerning the nature of the trajectories at the boundary. In order to address

these questions precisely we begin by defining the set

$$U^{\delta} = \left\{ (x, I, \gamma) \,|\, |x - \tilde{x}_0(I)| \leq \delta,\ \tilde{I}_1 \leq I \leq \tilde{I}_2 \right\}, \tag{3.1}$$

where

$$I_1 \leq \tilde{I}_1 < \tilde{I}_2 \leq I_2.$$

If $I_1 = \tilde{I}_1$ and $I_2 = \tilde{I}_2$ then clearly U^{δ} is a neighborhood of $\mathcal{M}$. However, for technical reasons (to be discussed shortly) we may need to slightly restrict the range of I values in discussing the perturbed manifolds and it is for this reason that the I interval in the definition of U^{δ} has been restricted (note: It will turn out that $\tilde{I}_1$ can be chosen arbitrarily close to I_1 and $\tilde{I}_2$ can be chosen arbitrarily close to I_2). The set U^{δ} will be useful in characterizing the nature of trajectories near the invariant manifolds. For the unperturbed system, we define the local stable and unstable manifolds of $\mathcal{M}$ as

$$W^{s}_{loc}(\mathcal{M}) \equiv W^{s}(\mathcal{M}) \cap U^{\delta}, \quad W^{u}_{loc}(\mathcal{M}) \equiv W^{u}(\mathcal{M}) \cap U^{\delta}.$$

We now state the persistence theorem.

Theorem 3.1. There exists $\epsilon_0 > 0$ sufficiently small such that for $0 < \epsilon < \epsilon_0$ $\mathcal{M}$ persists as a C^r, locally invariant two-dimensional normally hyperbolic manifold with boundary, which we denote by $\mathcal{M}_\epsilon$, having the following properties.

(1) $\mathcal{M}_\epsilon$ is C^r in ϵ and μ.

(2) $\mathcal{M}_\epsilon$ is C^r ϵ-close to $\mathcal{M}$ and can be represented as a graph over $\mathcal{M}$ as

$$\mathcal{M}_\epsilon = \left\{ (x, I, \gamma) | x = \tilde{x}_\epsilon(I, \gamma, \mu) = \tilde{x}_0(I) + \epsilon \tilde{x}_1(I, \gamma, \mu) + \mathcal{O}(\epsilon^2),\ \tilde{I}_1 \leq I \leq \tilde{I}_2 \right\}. \tag{3.2}$$

Moreover, there exists δ_0 sufficiently small (depending on ϵ) such that for $0 < \delta < \delta_0$ there exists locally invariant manifolds in U^{δ}, denoted $W^{s}_{loc}(\mathcal{M}_\epsilon), W^{u}_{loc}(\mathcal{M}_\epsilon)$, having the following properties.

(3) $W^{s}_{loc}(\mathcal{M}_\epsilon)$ and $W^{u}_{loc}(\mathcal{M}_\epsilon)$ are C^r in ϵ and μ.

(4) $W^{s}_{loc}(\mathcal{M}_\epsilon) \cap W^{u}_{loc}(\mathcal{M}_\epsilon) = \mathcal{M}_\epsilon$.

(5) $W^{s}_{loc}(\mathcal{M}_\epsilon)$ (resp. $W^{u}_{loc}(\mathcal{M}_\epsilon)$) is a graph over $W^{s}_{loc}(\mathcal{M})$ (resp. $W^{u}_{loc}(\mathcal{M})$) and is C^r ϵ-close to $W^{s}_{loc}(\mathcal{M})$ (resp. $W^{u}_{loc}(\mathcal{M})$).

(6) Let $(x^{s}_\epsilon(t), I^{s}_\epsilon(t), \gamma^{s}_\epsilon(t))$ (resp. $(x^{u}_\epsilon(t), I^{u}_\epsilon(t), \gamma^{u}_\epsilon(t))$) denote a trajectory that is in $W^{s}_{loc}(\mathcal{M}_\epsilon)$ (resp. $W^{u}_{loc}(\mathcal{M}_\epsilon)$) at $t = 0$. Then as $t \to +\infty$ ($t \to -\infty$), either

(a) $(x^{s}_\epsilon(t), I^{s}_\epsilon(t), \gamma^{s}_\epsilon(t))$ (resp. $(x^{u}_\epsilon(t), I^{u}_\epsilon(t), \gamma^{u}_\epsilon(t))$ crosses ∂U^{δ}

or

(b) $|x^{s}_\epsilon(t) - \tilde{x}_\epsilon(I, \gamma, \mu)| \to 0$ (resp. $|x^{u}_\epsilon(t) - \tilde{x}_\epsilon(I, \gamma, \mu)| \to 0$).

We refer to $W^{s}_{loc}(\mathcal{M}_\epsilon)$ and $W^{u}_{loc}(\mathcal{M}_\epsilon)$ as the local stable and unstable manifolds of $\mathcal{M}_\epsilon$, respectively.

Proof. The proof follows from the invariant manifold theory developed by Fenichel in ref. [12]. Fenichel's theory is adapted to systems of the type $(1.1)_\epsilon$ in [7], where the details of the proof of this theorem can be found. □

We make the following remarks concerning the consequences and implications of this theorem.

Remark 1. The term *locally invariant* means that trajectories with initial conditions on $\mathcal{M}_\epsilon$ may leave $\mathcal{M}_\epsilon$, however they may do so only by crossing the boundary of $\mathcal{M}_\epsilon$. In proving the persistence of $\mathcal{M}$ under perturbation it is necessary to know the stability properties of trajectories in $\mathcal{M}$ on semi-infinite time intervals. Technically, this control is accomplished by modifying the perturbed vector field $(1.1)_\epsilon$ in an arbitrarily small neighborhood of the boundary of $\mathcal{M}$ by using C^∞ "bump functions"; this procedure is thoroughly explained in [7] and [17]. The perturbed manifold is then constructed as a graph over the unperturbed manifold by using the graph transform technique. This is the reason why the range of I values for which $\mathcal{M}_\epsilon$ exists in the perturbed vector field $(1.1)_\epsilon$ may need to be slightly decreased.

Remark 2. The fact that $\mathcal{M}_\epsilon$ is C^r $(r \geq 2)$ in μ and ϵ allows us to Taylor expand the manifold in powers of μ and ϵ. This will be important because we may need to explicitly compute the $\mathcal{O}(\epsilon)$ term in the expansion of $\tilde{x}_\epsilon(I, \gamma, \mu)$ given in (3.2). We now explain how this may be done.

Differentiation of $\tilde{x}_\epsilon(I, \gamma, \mu)$ along the perturbed vector field $(1.1)_\epsilon$ gives a quasilinear partial differential equation that $\tilde{x}_\epsilon(I, \gamma, \mu)$ must satisfy. This equation is given by

$$\begin{aligned} &J\mathrm{D}_x H(\tilde{x}_\epsilon, I) + \epsilon g^x(\tilde{x}_\epsilon, I, \gamma, \mu\epsilon) \\ &\quad = \epsilon(\mathrm{D}_I \tilde{x}_\epsilon)\, g^I(\tilde{x}_\epsilon, I, \gamma, \mu, \epsilon) + (\mathrm{D}_\gamma \tilde{x}_\epsilon)(\mathrm{D}_I H(\tilde{x}_\epsilon, I) + \epsilon g^\gamma(\tilde{x}_\epsilon, I, \gamma, \mu, \epsilon)). \end{aligned} \tag{3.3}$$

We can differentiate (3.3) with respect to ϵ and obtain equations that the derivatives of $\tilde{x}_\epsilon(I, \gamma, \mu)$ must satisfy. In this way we find that $\tilde{x}_1(I, \gamma, \mu)$ must satisfy the ordinary differential equation:

$$\begin{aligned} &-(\mathrm{D}_\gamma \tilde{x}_1)\mathrm{D}_I H(\tilde{x}_0(I), I) + J\mathrm{D}_x^2 H(\tilde{x}_0(I), I)\tilde{x}_1 \\ &\quad = (\mathrm{D}_I \tilde{x}_0(I)) g^I(\tilde{x}_0(I), I, \gamma, \mu, 0) - g^x(\tilde{x}_0(I), I, \gamma, \mu, 0). \end{aligned} \tag{3.4}$$

We want to point out that in one special circumstance (indeed, the situation that will be most important to us) the solution of (3.4) can immediately be written down, namely, at resonance. For at resonance, i.e., $I = I^r$, we have $\mathrm{D}_I H(\tilde{x}_0(I^r), I^r) = 0$ so that (3.4) reduces to an algebraic equation with solution

$$\tilde{x}_1 = \left(J\mathrm{D}_x^2 H(\tilde{x}_0(I^r), I^r)\right)^{-1}\left[\mathrm{D}_I \tilde{x}_0(I^r)\, g^I(\tilde{x}_0(I^r), I^r, \gamma, \mu, 0) - g^x(\tilde{x}_0(I^r), I^r, \gamma, \mu, 0)\right]. \tag{3.5}$$

It is also easy to find an expression for $\mathrm{D}_I \tilde{x}_0(I)$ by implicitly differentiating the equation

$$\mathrm{D}_x H(\tilde{x}_0(I), I) = 0.$$

This simple calculation gives

$$\mathrm{D}_I \tilde{x}_0(I) = -\left(\mathrm{D}_x^2 H(\tilde{x}_0(I), I)\right)^{-1}(\mathrm{D}_I \mathrm{D}_x H(\tilde{x}_0(I), I)). \tag{3.6}$$

Eqs. (3.5) and (3.6) will be useful later on. We remark that invertibility of $\mathrm{D}_x^2 H(\tilde{x}_0(I), I)^{-1}$ follows from the hyperbolicity of $\tilde{x}_0(I)$.

Remark 3. We define the global stable and unstable manifolds of $\mathcal{M}_\epsilon$, denoted $W^s(\mathcal{M}_\epsilon)$ and $W^u(\mathcal{M}_\epsilon)$, respectively, as follows. Let $\phi_t(\cdot)$ denote the flow generated by $(1.1)_\epsilon$, then we define

$$W^s(\mathcal{M}_\epsilon) = \bigcup_{t\le 0} \phi_t\big(W^s_{\mathrm{loc}}(\mathcal{M}_\epsilon) \cap U^\delta\big), \quad W^u(\mathcal{M}_\epsilon) = \bigcup_{t\ge 0} \phi_t\big(W^u_{\mathrm{loc}}(\mathcal{M}_\epsilon) \cap U^\delta\big). \tag{3.7}$$

Remark 4. The phrase *stable manifold of an invariant set* typically means the manifold of trajectories that approach the invariant set as $t \to +\infty$. However, our definition has a slightly different meaning that is peculiar to our invariant set, i.e., $\mathcal{M}_\epsilon$, having a boundary. This is characterized in terms of the alternatives (a) and (b) of part (vi) of theorem 3.1; similarly for the *unstable manifold of an invariant set*.

3.2. The dynamics on $\mathcal{M}_\epsilon$

We now want to discuss the dynamics on $\mathcal{M}_\epsilon$ under the perturbed vector field $(1.1)_\epsilon$ for which we will introduce the (x, h, γ) coordinates. The perturbed vector field restricted to $\mathcal{M}_\epsilon$ is given by

$$\dot I = \epsilon g^I(\tilde x_\epsilon(I,\gamma,\mu), I, \gamma, \mu, \epsilon), \quad \dot\gamma = \mathrm{D}_I H(\tilde x_\epsilon(I,\gamma,\mu), I) + \epsilon g^\gamma(\tilde x_\epsilon(I,\gamma,\mu), I, \gamma, \mu, \epsilon). \tag{3.8}$$

Taylor expanding (3.8) in powers of ϵ gives

$$\begin{aligned}
\dot I &= \epsilon g^I(\tilde x_0(I), I, \gamma, \mu, 0) + \epsilon^2\big[\langle \mathrm{D}_x g^I(\tilde x_0(I), I, \gamma, \mu, 0), \tilde x_1(I,\gamma)\rangle + \mathrm{D}_\epsilon g^I(\tilde x_0(I), I, \gamma, \mu, 0)\big] \\
&\quad + \mathcal{O}(\epsilon^3), \\
\dot\gamma &= \mathrm{D}_I H(\tilde x_0(I), I) + \epsilon\big[\langle \mathrm{D}_x(\mathrm{D}_I H(\tilde x_0(I), I)), \tilde x_1(I,\gamma)\rangle + g^\gamma(\tilde x_0(I), I, \gamma, \mu, 0)\big] + \mathcal{O}(\epsilon^2),
\end{aligned} \tag{3.9}$$

where $\langle \cdot, \cdot \rangle$ represents the usual Euclidean inner product.

We want to study the dynamics of (3.9) near the resonance $I = I^r$. For this purpose we will change variables in order to derive a simpler equation that describes the dynamics in a neighborhood of the resonance. Substituting

$$I = I^r + \sqrt{\epsilon}\, h, \quad \gamma = \gamma \tag{3.10}$$

into (3.9), Taylor expanding in I in powers of $\sqrt{\epsilon}$, and rescaling time by letting $\tau = \sqrt{\epsilon}\, t$ gives the equations

$$\begin{aligned}
h' &= g^I + \sqrt{\epsilon}\, G(h,\gamma,\mu) + \mathcal{O}(\epsilon), \\
\gamma' &= \big(\langle \mathrm{D}_x(\mathrm{D}_I H), \mathrm{D}_I \tilde x_0\rangle + \mathrm{D}_I^2 H\big)h + \sqrt{\epsilon}\, F(h,\gamma,\mu) + \mathcal{O}(\epsilon),
\end{aligned} \tag{3.11}$$

where the prime denotes differentiation with respect to the rescaled time τ,

$$G(h,\gamma,\mu) = \big(\langle \mathrm{D}_x g^I, \mathrm{D}_I \tilde x_0\rangle + \mathrm{D}_I g^I\big)h$$

and

$$\begin{aligned}
F(h,\gamma,\mu) = \tfrac{1}{2}\big[&\langle (\mathrm{D}_x(\mathrm{D}_x \mathrm{D}_I H))\mathrm{D}_I \tilde x_0, \mathrm{D}_I \tilde x_0\rangle + \langle \mathrm{D}_x(\mathrm{D}_I H), \mathrm{D}_I^2 \tilde x_0\rangle + 2\langle \mathrm{D}_x(\mathrm{D}_I^2 H), \mathrm{D}_I \tilde x_0\rangle + \mathrm{D}_I^3 H\big]h^2 \\
&+ \langle \mathrm{D}_x(\mathrm{D}_I H), \tilde x_1\rangle + g^\gamma
\end{aligned}$$

and where all functions are evaluated at $\tilde{x}_0(I) = \tilde{x}_0(I^r)$, $I = I^r$, γ, μ, and $\epsilon = 0$. The important advantage gained in localizing the equations near the resonance is that at $\epsilon = 0$ (3.11) is a one-degree-of-freedom (hence, integrable) Hamiltonian system given by

$$h' = g^I = -D_\gamma \mathscr{H}, \quad \gamma' = \left(\langle D_x(D_I H), D_I \tilde{x}_0\rangle + D_I^2 H\right)h = D_h \mathscr{H}, \tag{3.12}$$

where

$$\mathscr{H}(h,\gamma) = \left(\langle D_x(D_I H), D_I \tilde{x}_0\rangle + D_I^2 H\right)\tfrac{1}{2}h^2 - \int_{\gamma_0}^{\gamma} g^I \, d\bar{\gamma}$$

is the Hamiltonian function. The integrable Hamiltonian structure at leading order is typical near resonances, see e.g., [13, 14, 16], and is extremely useful for understanding the qualitative (as well as the quantitative) structure of the dynamics near the resonance on $\mathscr{M}_\epsilon$. The $\sqrt{\epsilon}$ dependence in the change of variables given in (3.10) is a consequence of the fact that $(\langle D_x(D_I H(\tilde{x}_0(I^r), I^r)), D_I \tilde{x}_0\rangle + D_I^2 H(\tilde{x}_0(I^r), I^r)) \neq 0$. If this is violated then a scaling with a different fractional power of ϵ is required; see [14] for details.

From the point of view of the dynamics on $\mathscr{M}_\epsilon$ we will henceforth only be interested in the dynamics in an $\mathscr{O}(\sqrt{\epsilon})$ neighborhood of the resonance $I = I^r$ and it will be useful to introduce some notation that emphasizes this fact. We will be studying the dynamics in the annulus centered at $I = I^r$ denoted as follows:

$$\mathscr{A}_\epsilon \equiv \left\{(x,h,\gamma) | x = \tilde{x}_\epsilon(I^r + \sqrt{\epsilon}\, h, \gamma, \mu), |h| < C\right\},$$

where $C > 0$ is some constant. The constant C is chosen sufficiently large so that the annulus contains the unperturbed homoclinic orbit. *It is important to note that in the h–γ coordinates the resonance zone (i.e., the annulus $\mathscr{A}_\epsilon$) is of $\mathscr{O}(1)$ width.* The three dimensional stable and unstable manifolds of $\mathscr{A}_\epsilon$, denoted $W^s(\mathscr{A}_\epsilon)$ and $W^u(\mathscr{A}_\epsilon)$, respectively, are subsets of $W^s(\mathscr{M}_\epsilon)$ and $W^u(\mathscr{M}_\epsilon)$, respectively, that are obtained by restricting the I values appropriately.

From the point of view of perturbation theory, we will want to compare the dynamics in $\mathscr{A}_\epsilon$ with the unperturbed dynamics in the same region on $\mathscr{M}_\epsilon$ (this is described precisely in section 3.3). For this reason we define the "unperturbed annulus" as

$$\mathscr{A}_0 \equiv \left\{(x,h,\gamma) | x = \tilde{x}_0(I^r), |h| < C\right\},$$

with its three dimensional stable and unstable manifolds $W^s(\mathscr{A}_0)$ and $W^u(\mathscr{A}_0)$ that coincide along a branch.

We make the following assumption on the structure of the integrable Hamiltonian system (3.12).

Assumption 3. For $\mu = \mu_0$ there exists $\gamma_c(\mu_0)$ and $\gamma_s(\mu_0)$ such that $q_0 = (h,\gamma) = (0, \gamma_s(\mu_0))$ is a hyperbolic saddle type fixed point of (3.12) and $p_0 = (h,\gamma) = (0, \gamma_c(\mu_0))$ is a center type fixed point of (3.12). Moreover, q_0 is connected to itself by a homoclinic orbit and p_0 is the only fixed point inside this homoclinic orbit, see fig. 3.1.

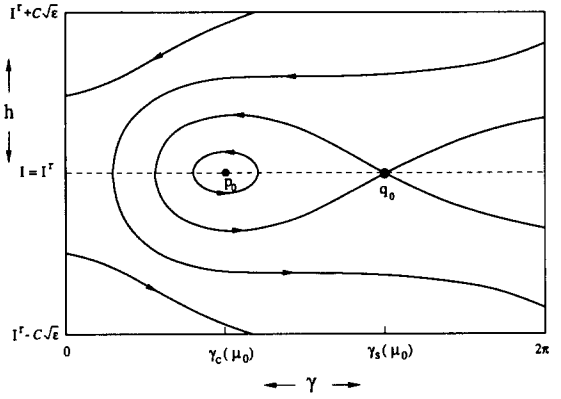

Fig. 3.1. The dynamics associated with the leading order Hamiltonian vector field restricted to $\mathcal{M}_\epsilon$ described in assumption 3.

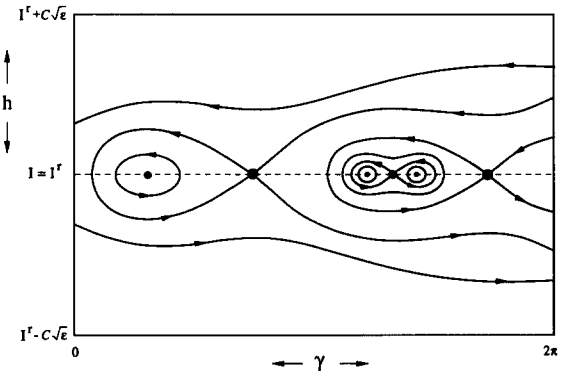

Fig. 3.2. Other possibilities for dynamics near the resonance for the leading order Hamiltonian vector field restricted to $\mathcal{M}_\epsilon$.

We make the following remarks concerning assumption 3 and its consequences.

Remark 1. It is possible for there to be more than one saddle-center pair with a homoclinic connection on the resonance such as we illustrate in fig. 3.2. In this case our methods can be applied to each pair separately.

Remark 2. Since the matrix associated with the linearization of the vector field about p_0 and q_0 is invertible, these fixed points will typically exist for an open set of parameter values that contains μ_0. A typical way for these fixed points to disappear as the parameters are varied is for them to coalesce in a Hamiltonian saddle-node bifurcation. However, if more fixed points and homoclinic (or heteroclinic) orbits exist, then other scenarios are possible.

Remark 3. An equation for the separatrix curve can be easily obtained from the Hamiltonian (3.12) by using the fact that the "energy" of the level set of $\mathcal{H}$ that defines the separatrix is equal to the energy of the saddle point q_0, i.e.,

$$\mathcal{H}(h,\gamma,\mu_0) - \mathcal{H}(0,\gamma_s(\mu_0),\mu_0) = 0. \tag{3.13}$$

Remark 4. Since (3.12) is an integrable Hamiltonian system obviously the $\mathcal{O}(\sqrt{\epsilon})$ terms in (3.11) may have a dramatic effect on the phase portrait. In particular, we are interested in the effect on the fixed points q_0 and p_0 and the homoclinic orbit connecting q_0 as well as the dynamics inside the region bounded by this homoclinic orbit. The following two lemmas address these issues.

Lemma 3.2. For ϵ sufficiently small,

(1) q_0 remains a hyperbolic fixed point of saddle stability type, denoted q_ϵ, for (3.11).

(2) If

$$\left(\left\langle D_x(D_I H), D_\gamma \tilde{x}_1\right\rangle + D_\gamma g^\gamma + D_I g^I + \left\langle D_x g^I, D_I \tilde{x}_0\right\rangle\right) < 0 \tag{3.14}$$

inside the homoclinic orbit connecting q_0, then p_0 becomes a hyperbolic sink, denoted p_ϵ, for (3.11) and the homoclinic orbit breaks with a branch of the unstable manifold of q_ϵ falling into the sink, p_ϵ, as shown in fig. 3.3.

(3) p_0 is $\mathcal{O}(\sqrt{\epsilon})$ close to p_ϵ and q_0 is $\mathcal{O}(\sqrt{\epsilon})$ close to q_ϵ.

Proof. Part (1) follows from the persistence of hyperbolic fixed points. Part (2) uses the fact that the quantity

$$\left(\left\langle D_x D_I H, D_\gamma \tilde{x}_1\right\rangle + D_\gamma g^\gamma + D_I g^I + \left\langle D_x g^I, D_I \tilde{x}_0\right\rangle\right)$$

is just $D_\gamma F + D_h G$, which is the leading order term of the trace of the linearization of (3.11). Then a routine analysis using Bendixson's criteria and some simple phase plane techniques gives the result, see, e.g., [13–16] for details. Part (3) follows from a simple application of the implicit function theorem. □

The area enclosed by the homoclinic orbit connecting q_0 in (3.12) is a good approximation to the domain of attraction of the sink p_ϵ. The following lemma makes this more precise.

Lemma 3.3. Suppose condition (3.14) of Lemma 3.2 holds and let p denote any point in the region bounded by the unperturbed homoclinic orbit that is an $\mathcal{O}(1)$ distance (in the h–γ coordinates) from the homoclinic orbit. Then, for ϵ sufficiently small, under the perturbed dynamics (i.e., (3.11)) the trajectory through p approaches p_ϵ asymptotically as $\tau \to \infty$.

Proof. Under condition (3.14) of lemma 3.2, the stable manifold of q_ϵ forms the boundary of the basin of attraction for p_ϵ. A standard planar Melnikov analysis of (3.11) shows that the stable and unstable manifolds of q_ϵ split by an $\mathcal{O}(\sqrt{\epsilon})$ amount (in the h–γ coordinates) in a tubular neighborhood of the unperturbed homoclinic orbit *excluding a small fixed neighborhood of* q_0, hence the result follows; see fig. 3.4. □

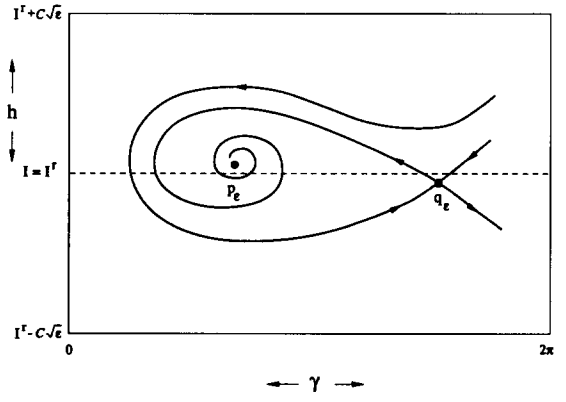

Fig. 3.3. The dynamics near the resonance of the vector field restricted to $\mathcal{M}_\epsilon$ under the conditions given in lemma 3.2.

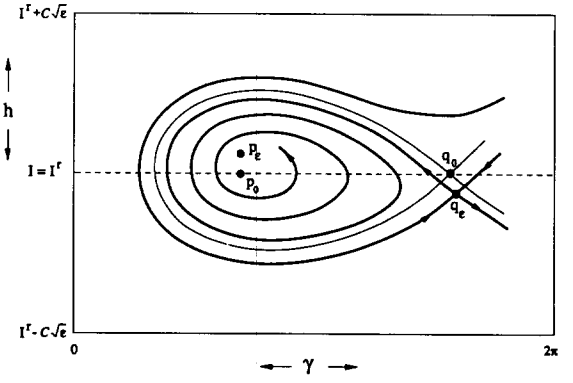

Fig. 3.4. The basic of attraction of p_ϵ and its comparison with the unperturbed structure (for the "slow time" system) near the resonance.

3.3. The fibering of $W^s(\mathscr{A}_\epsilon)$ and $W^u(\mathscr{A}_\epsilon)$: The singular perturbation nature

Theorem 3.1 provides a description of the asymptotic behavior of orbits in $W^s(\mathscr{A}_\epsilon)$ and $W^u(\mathscr{A}_\epsilon)$ and lemmas 3.2 and 3.3 describe the dynamics on $\mathscr{A}_\epsilon$. Now we want to "tie together" these results. In particular, we want to characterize orbits in $W^s(\mathscr{A}_\epsilon)$ and $W^u(\mathscr{A}_\epsilon)$ in terms of the orbits to which they asymptote in $\mathscr{A}_\epsilon$. We will see that this is a delicate problem in *singular perturbation theory* since the dynamics on $\mathscr{A}_\epsilon$ are "slow" compared to the "fast" dynamics transverse to $\mathscr{A}_\epsilon$. Indeed, the dynamics near the resonance as described in section 3.2 are created entirely by the perturbation. For these issues the (x, h, γ) coordinates will be most appropriate.

More insight into these questions can be obtained by directly examining the equations of motion. As described in section 3.2, the dynamics on $\mathscr{A}_\epsilon$ can be studied by restricting the vector field to $\mathscr{A}_\epsilon$ and introducing coordinates localized near the resonance as in (3.10). We rewrite these equations *without rescaling time*:

$$\begin{aligned}
\dot{h} &= \sqrt{\epsilon}\, g^I + \epsilon G(h, \gamma, \mu) + \mathscr{O}(\epsilon^{3/2}), \\
\dot{\gamma} &= \sqrt{\epsilon}\left(\langle \mathrm{D}_x(\mathrm{D}_I H), \mathrm{D}_I \tilde{x}_0\rangle + \mathrm{D}_I^2 H\right)h + \epsilon F(h, \gamma, \mu) + \mathscr{O}(\epsilon^{3/2}).
\end{aligned} \tag{3.15}$$

Note that for $\epsilon = 0$ (3.15) reduces to

$$\dot{h} = 0, \quad \dot{\gamma} = 0. \tag{3.16}$$

Of course, the dynamics in the full phase space are described by

$$\begin{aligned}
\dot{x} &= J\mathrm{D}_x H(x, I^r) + \sqrt{\epsilon}\,\mathrm{D}_I(J\mathrm{D}_x H(x, I^r))h + \tfrac{1}{2}\epsilon \mathrm{D}_I^2(J\mathrm{D}_x H(x, I^r))h^2 \\
&\quad + \epsilon g^x(x, I^r, \gamma, \mu, 0) + \mathscr{O}(\epsilon^{3/2}), \\
\dot{h} &= \sqrt{\epsilon}\, g^I(x, I^r, \gamma, \mu, 0) + \epsilon \mathrm{D}_I g^I(x, I^r, \gamma, \mu, 0)h + \mathscr{O}(\epsilon^{3/2}), \\
\dot{\gamma} &= \mathrm{D}_I H(x, I^r) + \sqrt{\epsilon}\,\mathrm{D}_I^2 H(x, I^r)h + \tfrac{1}{2}\epsilon \mathrm{D}_I^3 H(x, I^r)h^2 + \epsilon g^\gamma(x, I^r, \gamma, \mu, 0) + \mathscr{O}(\epsilon^{3/2}).
\end{aligned} \tag{3.17}$$

For $\epsilon = 0$, (3.17) reduces to

$$\dot{x} = J\mathrm{D}_x H(x, I^r), \quad \dot{h} = 0, \quad \dot{\gamma} = \mathrm{D}_I H(x, I^r). \tag{3.18}$$

There are several features that we want to point out concerning the above sets of equations.

(1) From (3.16), we see that at $\epsilon = 0$ the neighborhood of the resonance on $\mathscr{A}$ *in the variables scaled as in* (3.10) consists entirely of fixed points. We can think of the change of variables in (3.10) as "blowing up" the circle of fixed points into an annulus of fixed points centered at $I = I^r$.

(2) For ϵ small, but nonzero, we see that, roughly speaking, the character of the dynamics in the x variables is not altered much under the influence of the perturbation. Theorem 3.1 makes this more precise. However, the dynamics on the annulus are radically different. Indeed, for $\epsilon = 0$ there are no dynamics on the annulus (it consists entirely of fixed points), whereas for ϵ small the typical resonance structure is created as was revealed through a study of (3.15) under the rescaled time or, *slow time*, $\tau = \sqrt{\epsilon}\, t$.

As we mentioned earlier, we want to relate the asymptotic behavior of trajectories in the stable and unstable manifolds of the annulus to trajectories in the annulus for ϵ small, but nonzero. From eqs. (3.15)–(3.18) as well as the discussion following these equations it should be clear that this is a singular perturbation problem; however, since we will require infinite time results, classical singular perturbation approaches will be of limited use. Rather, we will see that the problem is most naturally addressed from the geometrical, dynamical systems viewpoint that utilizes the *fibering* of the stable and unstable manifolds by submanifolds consisting of initial conditions of trajectories that have the same "asymptotic phase".

We will set the stage for this by noting that one can interpret these unperturbed heteroclinic orbits connecting the fixed points in the annulus in a slightly different way. Clearly, all points on a heteroclinic orbit approach the same fixed point in the annulus asymptotically in forward time as well as the same fixed point in the annulus asymptotically in backward time (of course, the forward and backward time limit points are, in general, different). Thus, we see that $W^s(\mathscr{A}_0)$ and $W^u(\mathscr{A}_0)$ can be viewed as the union of a two parameter family of curves (the two parameters are h and γ) that have the properties that points on the curves asymptotically approach the same orbit (which is just a fixed point) in $\mathscr{A}_0$. We say that $W^s(\mathscr{A}_0)$ and $W^u(\mathscr{A}_0)$ are *fibered* by an invariant family of curves made up of initial conditions of trajectories that asymptote to the same orbit in $\mathscr{A}_0$. In the case of the perturbed problem, if we could prove that these fibers, as well as the dynamical interpretation that all points on the fibers have the same "asymptotic phase", were persistent.

We want to emphasize and illustrate the most important points with the following example which really captures the essence of the problem.

Consider the perturbed equation

$$\dot{\xi} = -\epsilon\xi, \quad \dot{\eta} = -\eta \tag{3.19}_\epsilon$$

with solution

$$\xi(t) = \xi_0 e^{-\epsilon t}, \quad \eta(t) = \eta_0 e^{-t} \tag{3.20}_\epsilon$$

and the unperturbed equation

$$\dot{\xi} = 0, \quad \dot{\eta} = -\eta \tag{3.19}_0$$

with solution

$$\xi(t) = \xi_0, \quad \eta(t) = \eta_0 e^{-t}. \tag{3.20}_0$$

The ξ axis is clearly an invariant manifold of the perturbed and unperturbed problems with the remainder of the ξ–η plane being its stable manifold. The analogy of this example with our problem should be clear. The ξ axis is analogous to our annulus and the remainder of the ξ–η plane is analogous to the stable manifold of the annulus.

Our goal is to relate orbits in the stable manifold of the perturbed problem to orbits in the invariant manifold, i.e., the ξ axis. Moreover, we would like to do this in a perturbative manner. Namely, assuming that we know "everything" about trajectories in the unperturbed problem, show that the perturbed trajectories are "close" to the unperturbed trajectories. However, immediately we see that this approach

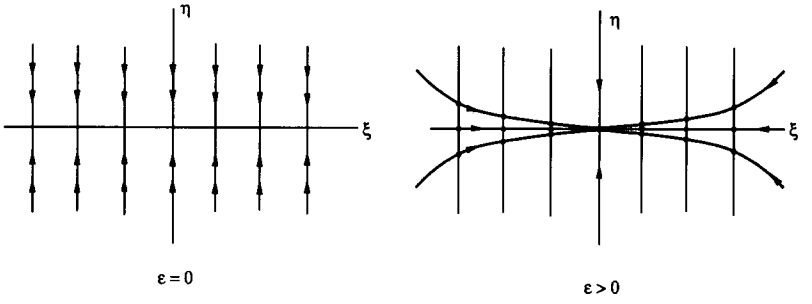

Fig. 3.5. Trajectories and fibers in an example problem.

may lead to serious difficulties. From $(3.20)_\epsilon$ and $(3.20)_0$ we see that the difference between trajectories in the perturbed and unperturbed problems is finite as $t \to \infty$. Moreover, the orbits of the perturbed vector field are given by the one-parameter family of parabolas $\{\eta = c\xi^{1/\epsilon} | c \in \mathbb{R}\}$; clearly, these orbits are singular at $\epsilon = 0$, we illustrate the trajectories of the perturbed and unperturbed systems in fig. 3.5. Therefore, it would seem that an approach based on a comparison of trajectories of the perturbed and unperturbed problems would not be successful. Let us consider the geometry more carefully and an alternate point of view.

In the unperturbed problem the vertical lines $\xi = \xi_0$ have the property that all points on them approach the point $(\xi_0, 0)$ asymptotically at the rate e^{-t}. Since the stable manifold of the ξ axis is the union of all such vertical lines, we can use these vertical lines to relate orbits in the stable manifold to orbits in the corresponding invariant manifold. If we refer to $(\xi_0, 0)$ as the *base point* of the vertical line $\xi = \xi_0$, then we see that trajectories with initial points on the vertical line $\xi = \xi_0$ asymptotically approach the trajectory in the invariant manifold through the base point of the vertical line. Thus the two-dimensional stable manifold is *fibered* by one-dimensional manifolds that have the property that trajectories which start in the fibers asymptotically approach the trajectory in the invariant manifold that starts at the base point of the respective fiber.

Now consider the perturbed problem; we know that the difference between trajectories of the perturbed and unperturbed problems becomes finite asymptotically. However, we want to examine the possibility that the fibering of the stable manifold, as well as the meaning of the fibers, will be stable under perturbations. The trajectory in the invariant manifold through the point $(\xi_0, 0)$ at $t = 0$ is given by $(\xi(t) = \xi_0 e^{-\epsilon t}, 0)$ and the trajectory in the stable manifold through the point (ξ_0, η_0) at $t = 0$ is given by $(\xi(t) = \xi_0 e^{-\epsilon t}, \eta(t) = \eta_0 e^{-t})$. From these two expressions we easily see that this trajectory in the stable manifold approaches the corresponding orbit in the invariant manifold asymptotically. In other words, the stable manifold in the perturbed problem is fibered by one-dimensional curves (manifolds) (i.e., it is the union of a set of one-dimensional curves) and these curves have the property that trajectories through them asymptotically approach the trajectories in the invariant manifold through the base point of the curve. Thus, for this simple example we conclude

(1) In the unperturbed problem the stable manifold is fibered by one-dimensional curves (actually, vertical lines) having the property that trajectories starting on the curves asymptotically approach the trajectory in the invariant manifold starting at the base point of the curve.

(2) This situation is stable under perturbations, i.e., the fibering of the stable manifold persists and the fibers perturb by $\mathcal{O}(\epsilon)$ (actually, in this simple example, the unperturbed and perturbed fibers are identical). Moreover, the interpretation of the fibers remains the same; namely, they represent curves of

initial conditions having the property that trajectories starting at points on the curves asymptotically approach the trajectory in the invariant manifold that starts at the base point of the curve.

This fibering of the stable manifold allows us to view some aspects of a singular perturbation problem as a problem in regular perturbation theory.

Fiberings of stable (and unstable) manifolds of this form have been studied in the context of dynamical systems theory for many years, however, they have not been used much in applications. Fenichel [17] was among the first to apply such ideas to problems arising in singular perturbation theory; in particular, to problems of the form of (3.17). Recall again the geometrical structure of this problem. At $\epsilon = 0$ the annulus $\mathscr{A}_0$ consists entirely of fixed points (under the "fast time" t). By hypothesis, each fixed point is connected to another (in general) fixed point by a heteroclinic connection. Thus the three dimensional stable and unstable manifolds of the annulus are fibered by a two parameter family (the parameters label a point on the annulus) of one dimensional fibers; clearly, the fibers at $\epsilon = 0$ are just the unperturbed heteroclinic orbits. Fenichel has proven that the fibers perturb smoothly in $\sqrt{\epsilon}$ and are $\mathscr{O}(\sqrt{\epsilon})$ close to the unperturbed fibers. This is stated more precisely in the following theorem that we formulate in the context of eq. (3.17).

Theorem 3.4. There exists $\delta_0 > 0$ and $\epsilon_0 > 0$ such that given any point $(\bar{h}, \bar{\gamma}) \in \mathscr{A}_\epsilon$ there exists a family of one-dimensional curves, called the *stable fibers*, that can be represented as graphs as follows:

$$x_2 = x_2\left(x_1; \bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon}\right), \quad h = h\left(x_1; \bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon}\right), \quad \gamma = \gamma\left(x_1; \bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon}\right)$$

where $x \equiv (x_1, x_2)$. The point $(\bar{h}, \bar{\gamma})$ is referred to as the *basepoint* of the fiber. These graphs are defined for any $0 < \delta \leq \delta_0$, $0 < \epsilon \leq \epsilon_0$ with $|x_1| \leq \delta$ and $\mu \in \mathbb{R}^p$. Moreover, these curves have the following properties.

(1) They are C^r in x_1 and C^{r-1} in $(\bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon})$.

(2) $x_2(\tilde{x}_{1\epsilon}(I^r + \sqrt{\epsilon}\bar{h}, \bar{\gamma}, \mu); \bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon}) = \tilde{x}_{2\epsilon}(I^r + \sqrt{\epsilon}\bar{h}, \bar{\gamma}, \mu)$, $h(\tilde{x}_{1\epsilon}(I^r + \sqrt{\epsilon}\bar{h}, \bar{\gamma}, \mu); \bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon}) = \bar{h}$, $\gamma(\tilde{x}_{1\epsilon}(I^r + \sqrt{\epsilon}\bar{h}, \bar{\gamma}, \mu); \bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon}) = \bar{\gamma}$,
where, recall, $\mathscr{A}_\epsilon$, with the I values suitably restricted, is the graph of $\tilde{x}_\epsilon(I, \gamma, \mu) \equiv (\tilde{x}_{1\epsilon}(I, \gamma, \mu), \tilde{x}_{2\epsilon}(I, \gamma, \mu))$ (cf. theorem 3.1).

(3) $W^s_{loc}(\mathscr{A}_\epsilon)$ is the union of all stable fibers with basepoints in $\mathscr{A}_\epsilon$.

(4) Let $(\bar{h}(t), \bar{\gamma}(t))$ be a trajectory in $\mathscr{A}_\epsilon$ satisfying $(\bar{h}(0), \bar{\gamma}(0)) = (\bar{h}, \bar{\gamma})$ and let $(x_1(t), x_2(t), h(t), \gamma(t))$ be a trajectory in $W^s_{loc}(\mathscr{A}_\epsilon)$ satisfying

$$x_2(0) = x_2\left(x_1(0); \bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon}\right), \quad h(0) = h\left(x_1(0); \bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon}\right), \quad \gamma(0) = \gamma\left(x_1(0); \bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon}\right),$$

i.e., the trajectory starts on the fiber with basepoint $(\bar{h}, \bar{\gamma})$, then

$$\left|(x(t), h(t), \gamma(t)) - \left[\tilde{x}_\epsilon\left(I^r + \sqrt{\epsilon}\bar{h}(t), \bar{\gamma}(t), \mu\right), \bar{h}(t), \bar{\gamma}(t)\right]\right| < C\,\mathrm{e}^{-\lambda t}$$

for all $t > 0$ and for some C, $\lambda > 0$ as long as $(\bar{h}(t), \bar{\gamma}(t)) \in \mathscr{A}_\epsilon$. In other words, trajectories starting on a

stable fiber asymptotically approach the trajectory in $\mathscr{M}_\epsilon$ that starts on the basepoint of the same fiber, as long as the trajectory through this basepoint remains in $\mathscr{M}_\epsilon$.

(5) The family of fibers form an invariant family in the sense that fibers map to fibers under the time t flow map. Analytically, this is expressed as follows. Suppose $(x_1(t), x_2(t), h(t), \gamma(t))$ is a trajectory satisfying

$$x_2(0) = x_2\big(x_1(0); \bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon}\big), \quad h(0) = h\big(x_1(0); \bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon}\big), \quad \gamma(0) = \gamma\big(x_1(0); \bar{h}, \bar{\gamma}, \mu, \sqrt{\epsilon}\big),$$

then

$$x_2(t) = x_2\big(x_1(t); \bar{h}(t), \bar{\gamma}(t), \mu, \sqrt{\epsilon}\big), \quad h(t) = h\big(x_1(t); \bar{h}(t), \bar{\gamma}(t), \mu, \sqrt{\epsilon}\big),$$
$$\gamma(t) = \gamma\big(x_1(t); \bar{h}(t), \bar{\gamma}(t), \mu, \sqrt{\epsilon}\big).$$

(6) At $\epsilon = 0$ the unperturbed fibers correspond to the unperturbed heteroclinic orbits. Hence the perturbed and unperturbed fibers are $C^r\sqrt{\epsilon}$-close.

Proof. This follows immediately from the results of Fenichel [12, 17]. □

We make the following remarks concerning this theorem.

Remark 1. An identical result follows for the fibering of $W^u(\mathscr{M}_\epsilon)$.

Remark 2. Quasilinear partial differential equations whose solutions are the fibers can be derived. These equations are analogous to those given following theorem 3.1. We will not require these for our calculations, however, the reader can find these equations in [17].

Finally, we want to state a result that will be important in the next section.

Proposition 3.5. $W^u_{loc}(p_\epsilon)$ is C^r ϵ-close to $W^u_{loc}(p_0)$.

Proof. This result follows from a slight, but straightforward, modification of the usual unstable manifold theorem, see [15] for details. A modification of the usual result is required since p_0 is not hyperbolic. □

At this stage it is appropriate to warn the reader that "closeness" is a concept that depends on the specific coordinate system under consideration. In particular, in this paper we are considering two coordinate systems; the x–I–γ coordinate system and the x–h–γ coordinate system. *Points that are $\mathscr{O}(\epsilon)$ close in the x–I–γ coordinates are $\mathscr{O}(\sqrt{\epsilon})$ close in the x–h–γ coordinates.* Thus, proposition 3.5 is a statement about the closeness of $W^u_{loc}(p_\epsilon)$ and $W^u_{loc}(p_0)$ in the x–I–γ coordinates. These issues will play an important role in section 4.1.

4. The existence of a homoclinic connection to p_ϵ

The geometric structure of the perturbed system described in section 3 largely follows from general results related to the perturbation of normally hyperbolic invariant sets and their fiberings by submani-

folds of initial conditions with the same asymptotic phase. We will use this geometrical structure to develop the methods that will enable us to prove our main result; namely, the existence of an orbit homoclinic to the fixed point p_ϵ in the resonance. The proof will consist of two steps. In the first step we show that $W^u(p_\epsilon) \subset W^s(\mathscr{A}_\epsilon)$. Roughly speaking, this will imply that a trajectory in $W^u(p_\epsilon)$ leaves a neighborhood of the annulus $\mathscr{A}_\epsilon$ and returns close to the annulus. In the second step we show that the trajectory actually asymptotes to p_ϵ as $t \to \infty$. The first step involves an approach which is similar to the higher dimensional Melnikov theory developed in [7]. The second step involves the fibers of $W^s(\mathscr{A}_\epsilon)$ and is required even though from the first step we can conclude that trajectories in $W^u(p_\epsilon)$ asymptote to an orbit in $\mathscr{A}_\epsilon$, this orbit need not be in the domain of attraction of p_ϵ. We begin with the first step.

4.1. $W^u(p_\epsilon) \subset W^s(\mathscr{A}_\epsilon)$: *The higher dimensional Melnikov theory*

The setting here is the higher dimensional Melnikov theory as developed in [7], for which the (x, I, γ) coordinates are most appropriate. Our goal is to determine whether or not $W^u(p_\epsilon)$ intersects $W^s(\mathscr{A}_\epsilon)$. In order to do this we will develop a perturbative measure of the distance between these manifolds, of which the leading order term is (up to a nonzero normalization factor) a higher dimensional generalization of the *Melnikov function*. Since the details of these techniques can be readily found in the literature (cf. [7, 15]), here we merely summarize the essential points and steps in the development of the theory that we shall need.

Step 1. For the unperturbed system, develop a parametrization of the unperturbed homoclinic manifold and set up a moving system of "homoclinic coordinates" on the homoclinic manifold.

A parametrization of the unperturbed heteroclinic orbits in $W^s(\mathscr{A}_0) \cap W^u(\mathscr{A}_0)$ was given in section 2.2 (eq. (2.7)). With respect to this parametrization, we can "pick out" $W^u(p_0)$ by fixing the I and γ_0 values in the expressions for the heteroclinic orbits that correspond to the coordinates of $W^u(p_0)$. In particular, we fix $I = I^r$ and $\gamma_0 = \int_{-\infty}^{0} D_I H(x^h(s, I^r), I^r)\,ds + \gamma_c(\mu_0)$. This choice for γ_0 gives $\gamma(-\infty, I^r, \gamma_0) = \gamma_c(\mu_0)$. The remaining coordinate in the parametrization of $W^u(p_0)$ is t_0, and varying t_0 moves us around $W^u(p_0)$.

Now at $\epsilon = 0$ the one-dimensional $W^u(p_0)$ lies in the three-dimensional $W^s(\mathscr{A}_0)$. We want to determine under what conditions this situation will hold for the perturbed problem. We will do this by developing a perturbative measure of the distance between the corresponding perturbed manifolds. This distance measure uses the global geometry of the underlying integrable structure. In particular, let $p = (x^h(-t_0, I^r), I^r, \gamma(-t_0, I^r, \gamma_0))$ denote a particular point on $W^u(p_0) \cap W^s(\mathscr{A}_0)$ and consider the three-dimensional hyperplane at p, denoted Π_p, spanned by the vectors $\underline{n}(p) \equiv (D_x H(p), D_I H(p) - D_I H(\tilde{x}_0(I^r), I^r), 0)$, $\hat{I}$, and $\hat{\gamma}$ where $\hat{I}$ and $\hat{\gamma}$ denote constant unit vectors in the I and γ directions, respectively, and $\underline{n}(p)$ denotes the vector normal to Γ at the point p. Varying $t_0 \in \mathbb{R}$ serves to move this hyperplane to all points of $W^u(p_0) \cap W^s(\mathscr{A}_0)$ and $W^u(p_0)$ intersects the hyperplane at each point transversely in a point and $W^s(\mathscr{A}_0)$ intersects the hyperplane transversely at each point in a two-dimensional surface, see fig. 4.1. We develop a measure of the distance between $W^u(p_\epsilon)$ and $W^s(\mathscr{A}_\epsilon)$ *at the point p in the coordinates defining* Π_p.

Step 2. The persistence of $W^u(p_\epsilon)$ and $W^s(\mathscr{A}_\epsilon)$ and the persistence of transversal intersections of these manifolds with Π_p.

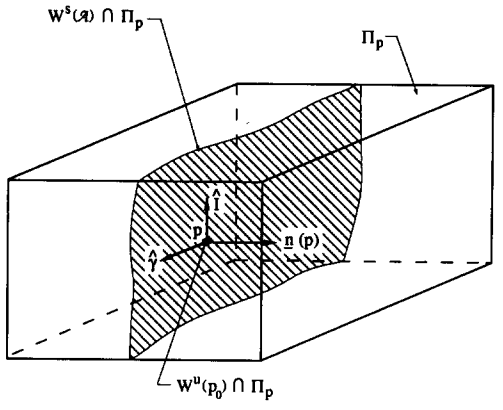

Fig. 4.1. Homoclinic coordinates in the unperturbed system.

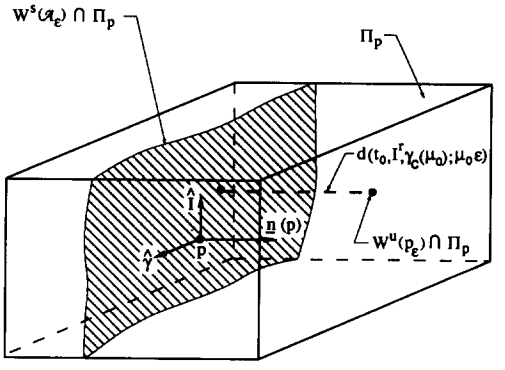

Fig. 4.2. The measurement of the splitting of the manifolds in the perturbed system.

For ϵ sufficiently small $W^u(p_0)$ and $W^s(\mathscr{A}_0)$ persist (denoted $W^u(p_\epsilon)$ and $W^s(\mathscr{A}_\epsilon)$, respectively). Moreover, $W^u_{loc}(p_0)$ and $W^s_{loc}(\mathscr{A}_0)$ are C^r ϵ-close to $W^u_{loc}(p_\epsilon)$ and $W^s_{loc}(\mathscr{A}_\epsilon)$, respectively. Using simple Gronwall estimates, it can easily be shown that outside a neighborhood of $\mathscr{A}_0$, trajectories in $W^u(p_0)$ are ϵ-close to trajectories in $W^u(p_\epsilon)$ for finite time intervals and trajectories in $W^s(\mathscr{A}_0)$ are ϵ-close to trajectories in $W^s(\mathscr{A}_\epsilon)$ for finite time intervals. Moreover, by persistence of transversal intersections, outside a neighborhood of $\mathscr{A}_0$, at each point $p = (x^h(-t_0, I^r), I^r, \gamma(-t_0, I^r, \gamma_0)) \in W^u(p_0) \cap W^s(\mathscr{A}_0)$, $W^u(p_\epsilon)$ intersects Π_p transversely in a point and $W^s(\mathscr{A}_\epsilon)$ intersects Π_p transversely in a two-dimensional surface, see fig. 4.2.

Step 3. Derive the Melnikov function and describe its geometric interpretation.

It can be shown (see [7, 15]) that the distance between $W^u(p_\epsilon)$ and $W^s(\mathscr{A}_\epsilon)$ at the point $p = (x^h(-t_0, I^r), I^r, \gamma(-t_0, I^r, \gamma_0))$, as measured in Π_p, is given by

$$d(t_0, I^r, \gamma_c(\mu_0); \mu_0, \epsilon) = \frac{\epsilon M(t_0, I^r, \gamma_c(\mu_0); \mu_0)}{\|\underline{n}(p)\|} + \mathscr{O}(\epsilon^2),$$

where

$$M(t_0, I^r, \gamma_c(\mu_0); \mu_0) = \int_{-\infty}^{+\infty} \big(\langle D_x H, g^x \rangle + (D_I H)(g^I)\big)\big(q_0(t - t_0, I^r, \gamma_c(\mu_0)), \mu_0, 0\big)\, dt$$
$$- (D_I H(x_0(I^r), I^r)) \int_{-\infty}^{+\infty} g^I\big(q_0(t - t_0, I^r, \gamma_c(\mu_0)), \mu_0, 0\big)\, dt, \tag{4.1}$$

with

$$q_0(t - t_0, I^r, \gamma_c(\mu_0))$$
$$= \Big(x^h(t - t_0, I^r), I^r, \gamma(t - t_0, I^r, \gamma_0) = \int_{-\infty}^{t - t_0} D_I H(x^h(s, I^r), I^r)\, ds + \gamma_c(\mu_0)\Big).$$

This is accomplished via the standard "Melnikov trick". Namely, a time-dependent Melnikov function is derived from trajectories of the perturbed vector field through the points on the perturbed manifolds in Π_p. By differentiation with respect to time, as well as use of the unperturbed vector field, a linear ordinary differential equation can be derived which is satisfied by the time dependent Melnikov function. This equation is readily solved and the initial point of the solution yields (4.1). Details can be found in [7]. The function $M(t_0, I^r, \gamma_c(\mu_0); \mu_0)$ is referred to as the Melnikov function. Note that the Melnikov function in (4.1) simplifies since $D_I H(x_0(I^r), I^r) = 0$ for the class of problems studied in this paper. We leave this term in, however, since it appears in the more general Melnikov function given in [7] and, also, we will refer to it in our calculations in section 6. The following characteristics of the Melnikov function are important.

(1) Note that if we make the change of variables $t \to t + t_0$ in (4.1) then t_0 disappears from the Melnikov function. Henceforth we assume this has been done and omit t_0 from the arguments of the Melnikov function. This has an important geometrical meaning that we next describe.

(2) If the one-dimensional $W^u(p_\epsilon)$ and the three-dimensional $W^s(\mathscr{A}_\epsilon)$ intersect at a single point, then, by uniqueness of solutions, they must coincide along a one-dimensional orbit. Since $W^u(p_\epsilon)$ is invariant, this implies that $W^u(p_\epsilon)$ is contained in $W^s(\mathscr{A}_\epsilon)$. This is the reason why t_0 can be eliminated from the Melnikov function since t_0 parametrizes points along a trajectory (this is explained in great detail in [7, 13, 15]). Hence, the measure of distance between the perturbed manifolds can be made at any point along the unperturbed trajectory.

We have the following theorem:

Theorem 4.1. Suppose that at $\mu = \mu_0$ (3.14) holds, and also

(1) $M(I^r, \gamma_c(\mu_0), \mu_0) = 0$,

(2) $(d/d\mu)M(I^r, \gamma_c(\mu_0), \mu_0)$ has rank 1,

then $W^u(p_\epsilon) \subset W^s(\mathscr{A}_\epsilon)$.

Proof. The proof can be found in [7] and [15]. □

Note that the Melnikov function is only a function of the parameters μ_0. This is not unexpected since we would expect that the one-dimensional $W^u(p_\epsilon)$ and the three-dimensional $W^s(\mathscr{A}_\epsilon)$ would generically intersect in a one-parameter family of vector fields, i.e., it is a codimension-one phenomenon.

4.2. $W^u(p_\epsilon) \cap W^s(p_\epsilon)$: *a homoclinic orbit to* p_ϵ

Our arguments in this section will apply to the rescaled equations (3.15) and (3.17). *In particular, all measures of "closeness" will be made with respect to the* x–h–γ *coordinates* (cf. the comment at the end of section 3). Suppose we have shown that $W^u(p_\epsilon) \subset W^s(\mathscr{A}_\epsilon)$, then it follows from proposition 3.5 that there exist points

$$\hat{p}_\epsilon \equiv \partial U^\delta \cap W^u_{loc}(p_\epsilon), \quad \hat{p}_0 \equiv \partial U^\delta \cap W^u_{loc}(p_0)$$

such that

$$|\hat{p}_\epsilon - \hat{p}_0| = \mathscr{O}(\sqrt{\epsilon}),$$

see fig. 4.3. Since $W^u(p_\epsilon) \subset W^s(\mathscr{A}_\epsilon)$, the trajectory through the point $\hat{p}_\epsilon$ at $t = 0$ will return to U^δ after

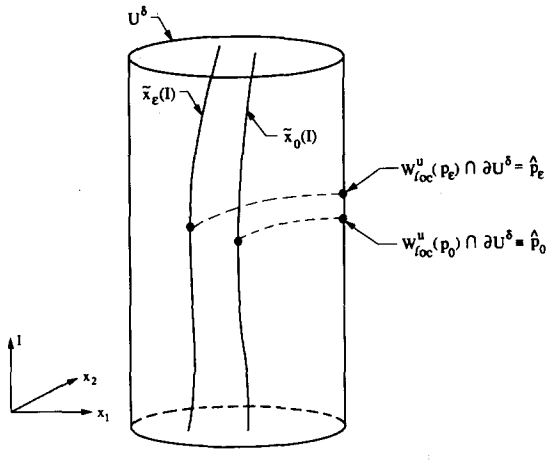

Fig. 4.3. The geometry associated with trajectories leaving a neighborhood of $\mathcal{M}_\epsilon$, i.e., the points $\hat{p}_\epsilon$ and $\hat{p}_0$, (with the γ coordinates suppressed).

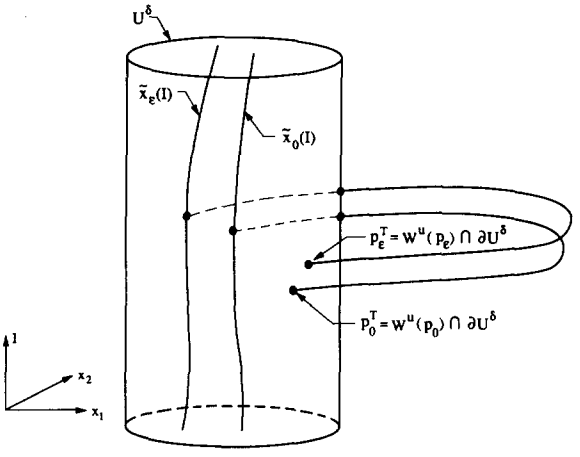

Fig. 4.4. The geometry associated with trajectories returning to a neighborhood of $\mathcal{M}_\epsilon$, i.e., the points p^T_ϵ and p^T_0, (with the γ coordinates suppressed).

some *finite* time of flight. We denote this point by

$$p^T_\epsilon \equiv \partial U^\delta \cap W^u(p_\epsilon).$$

From the unperturbed problem we also have a point

$$p^T_0 \equiv \partial U^\delta \cap W^u(p_0).$$

Since this time of flight from ∂U^δ to ∂U^δ is finite, it follows from simple Gronwall type estimates that

$$|p^T_0 - p^T_\epsilon| = \mathcal{O}(\sqrt{\epsilon}),$$

see fig. 4.4.

At this point we have shown the existence of an orbit that leaves a neighborhood of $\mathcal{M}_\epsilon$ and returns to a neighborhood of $\mathcal{M}_\epsilon$ by using perturbation theory of normally hyperbolic invariant manifolds coupled with a generalized Melnikov type analysis. It remains to show that the orbit approaches p_ϵ asymptotically as $t \to \infty$; to show this we must use the fibers. It follows from theorems 3.4 and 4.1 that the points p^T_0 and p^T_ϵ are on fibers. We denote the base points of these fibers by p^∞_0 and p^∞_ϵ, respectively, where, by theorem 4.1

$$|p^\infty_0 - p^\infty_\epsilon| = \mathcal{O}(\sqrt{\epsilon}).$$

Moreover, the h–γ coordinates of p^∞_0 are given by

$$p^\infty_0 \equiv (0, \gamma^\infty(\mu_0)) = (0, \gamma_c(\mu_0) + \Delta\gamma),$$

where $\Delta\gamma$ is given by (2.8). Now within $\mathcal{M}_\epsilon$ the domain of attraction of p_ϵ is approximated by the level set of the Hamiltonian connecting q_0 in the sense described in lemma 3.3. Moreover, in the h–γ coordinates, the area enclosed by this level set of the Hamiltonian is $\mathcal{O}(1)$. Hence, it follows from lemma 3.4 that for ϵ sufficiently small, if p^∞_0 is contained within the homoclinic loop connecting q_0 (and an

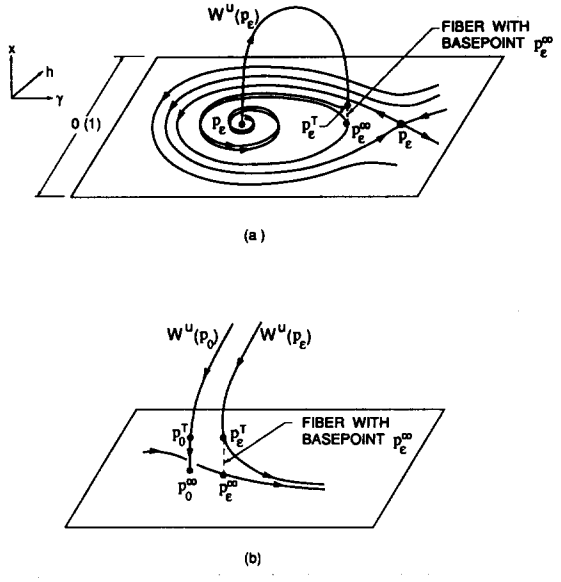

Fig. 4.5. (a) The geometry associated with the return of a trajectory to a neighborhood of $\mathcal{M}_\epsilon$ that is homoclinic to p_ϵ. (b) The perturbed and unperturbed fibers.

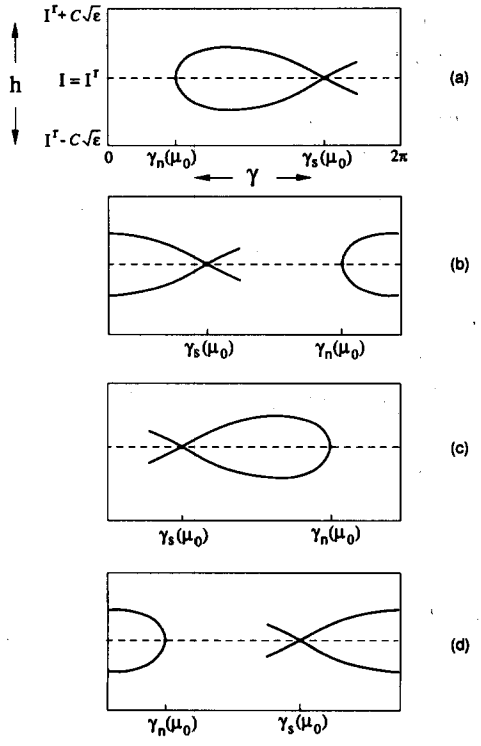

Fig. 4.6. Possible geometrical configurations for the unperturbed (on the "slow time" scale) homoclinic orbit on $\mathcal{M}_\epsilon$ that connects p_0.

$\mathcal{O}(1)$ distance away from the homoclinic loop) that p_ϵ^∞ will approach p_ϵ asymptotically as $t \to \infty$. It then follows from theorem 4.1 that the trajectory through p_ϵ^T will approach p_ϵ asymptotically as $t \to \infty$, see fig. 4.5.

Verifying this situation is straightforward. Recall from section 3.2 that the equation for the homoclinic loop connecting q_0 is given by

$$\mathcal{H}(h, \gamma, \mu) - \mathcal{H}(0, \gamma_s(\mu_0), \mu_0) = 0.$$

We will refer to the region bounded by this curve as "the fish". In terms of the angle values γ, the "tail" of the fish is at $\gamma = \gamma_s(\mu_0)$ and the "nose" of the fish is located at $\gamma = \gamma_n(\mu_0)$ where $\gamma_n(\mu_0)$ is found by solving the equation

$$\mathcal{H}(0, \gamma, \mu_0) - \mathcal{H}(0, \gamma_s(\mu_0), \mu_0) = 0.$$

In fig. 4.6 we illustrate the four general possibilities under our assumptions, which we refer to as cases (a)–(d). We summarize our results in the following theorem.

Theorem 4.2. Suppose that at $\mu = \mu_0$ assumption 3 and (3.14) holds, and also
(1) $M(I^r, \gamma_c(\mu_0), \mu_0) = 0$,
(2) $\mathrm{d}/\mathrm{d}\mu\, M(I^r, \gamma_c(\mu_0), \mu_0)$ has rank 1
with *one* of the following cases holding (refer to fig. 4.6 for an explanation of the different cases; all

angles are taken mod 2π)

(a) $\gamma_n(\mu_0) < \gamma^\infty(\mu_0) < \gamma_s(\mu_0)$,

(b) $0 \le \gamma^\infty(\mu_0) < \gamma_s(\mu_0)$ or $\gamma_n(\mu_0) < \gamma^\infty(\mu_0) \le 2\pi$,

(c) $\gamma_s(\mu_0) < \gamma^\infty(\mu_0) < \gamma_n(\mu_0)$,

(d) $0 \le \gamma^\infty(\mu_0) < \gamma_n(\mu_0)$ or $\gamma_s(\mu_0) < \gamma^\infty(\mu_0) \le 2\pi$,

then $(1.1)_\epsilon$ possesses a "simple" homoclinic orbit connecting p_ϵ. Moreover, if (1) does not hold, then there are no homoclinic orbits connecting p_ϵ and if (1) and the angle inequalities for the appropriate case (a)–(d) *do not* hold, then there are no "simple" homoclinic orbits connecting to p_ϵ.

By a "simple homoclinic orbit" we mean a homoclinic orbit that makes one excursion through a neighborhood of the annulus before connecting p_ϵ. Our methods say nothing about the existence of homoclinic orbits that may make several passes near the annulus before connecting p_ϵ.

Some remarks on the two different sets of coordinates

We now make some general observations on the role of the two sets of coordinates used throughout the analysis, the (x, I, γ) coordinates and the (x, h, γ) coordinates. As should be clear from the analysis, the (x, h, γ) coordinates serve to localize the system near the resonance on $\mathscr{A}_\epsilon$. This results in the restriction to the annulus $\mathscr{A}_\epsilon$, where the resonance structure is analyzed as well as the approach of trajectories to this resonance structure via the fibering of $W^s(\mathscr{A}_\epsilon)$. The (x, I, γ) coordinates are more global in nature and are most convenient for developing the geometry behind the higher dimensional Melnikov theory.

The most important, however, is the fact that for the purposes of applications, the equations are most naturally initially written down in the (x, I, γ) coordinate system. The reader will see this in section 6 in our analysis of the dynamics of the two mode truncation of the damped, driven nonlinear Schrödinger equation. Also, in applications one is often interested in understanding as completely as possible the global dynamics in the (x, I, γ) coordinates and it should be clear that the (x, h, γ) coordinates cannot be used for these purposes. An example of this can be found in the numerical study of this same two-mode truncation of the damped, driven nonlinear Schrödinger equation given in [9].

Also, the interplay between the x–I–γ coordinates and the x–h–γ coordinates is much like the situation encountered in the study of resonance bands in time-periodically perturbed one-degree-of-freedom Hamiltonian systems. In that setting one is often interested not only in the dynamics associated with a single resonance band, but in the behavior of many resonance bands and how they are interrelated. In this situation it is important to understand how the dynamics near a specific resonance band is affected by the structure of the "globally defined" vector field. This is analogous to understanding how equation (3.17) derives its structure from $(1.1)_\epsilon$.

5. Chaos: Silnikov's theorem

Perturbation of homoclinic orbits is a common mechanism for producing chaos in the sense of Smale horseshoes. Orbits homoclinic to fixed points in autonomous ordinary differential equations may or may not give rise to chaotic dynamics; additional conditions are needed concerning the stability properties of the fixed point (in particular, the dimensions of the stable and unstable manifolds) and geometrical properties of the intersection of the stable and unstable manifolds (e.g., the dimension of the intersec-

tion) that give rise to the homoclinic orbit. For the class of systems we are studying the following theorem of Silnikov [18] is relevant.

Theorem 5.1. Consider a four-dimensional, autonomous C^r, $r \geq 3$ ordinary differential equation having a hyperbolic fixed point that is connected to itself by a homoclinic orbit. Moreover, suppose that the matrix associated with the linearization at the fixed point has the form (in appropriate local coordinates)

$$\begin{pmatrix} -\rho & -\omega & 0 & 0 \\ \omega & -\rho & 0 & 0 \\ 0 & 0 & -\lambda & 0 \\ 0 & 0 & 0 & \nu \end{pmatrix}, \tag{5.1}$$

where

$$\rho, \lambda > 0, \quad \omega \neq 0, \quad \nu > 0 \quad \text{and} \quad \nu > \rho > 0, \quad \lambda \neq \nu. \tag{5.2}$$

Then a three-dimensional Poincaré map defined sufficiently close to the homoclinic orbit possesses an invariant Cantor set on which it is topologically conjugate to a subshift on a countable set of symbols.

The original proof can be found in [18]. A more geometrical proof can be found in [7]. We make the following comments concerning this result.

(1) It should be clear from the hypotheses of the theorem that the fixed point has a three-dimensional stable manifold and a one-dimensional unstable manifold.

(2) In our class of problems $-\rho \pm i\omega$ represent the eigenvalues of the vector field restricted to $\mathscr{A}_\epsilon$, so it follows that $-\rho$ and ω are $\mathscr{O}(\epsilon)$. The eigenvalues $-\lambda$ and ν represent linearized growth rates transverse to $\mathscr{A}_\epsilon$, so $-\lambda$ and ν are $\mathscr{O}(1)$. Hence, it follows immediately that $\nu > \rho > 0$. We need only verify that $\lambda \neq \nu$, however it should be clear that this will generically be the case.

(3) If the system is slightly perturbed in a way that the homoclinic connection is broken, then the Poincaré map has an invariant Cantor set on which it is topologically conjugate to a full shift on a finite number of symbols. Thus the chaos is stable under perturbation.

(4) If $\lambda > \nu$ then the orbits in the invariant Cantor set of the Poincaré map have one positive Lyapunov exponent and two negative Lyapunov exponents. If $\lambda < \nu$ then the orbits have two positive Lyapunov exponents and one negative Lyapunov exponents, see [7].

(5) The manifestation of the chaotic dynamics in a specific application requires an interpretation of the geometry underlying the construction of the chaotic invariant set. We will see an example of this in section 6.

(6) There are other possible mechanisms for chaotic dynamics in this class of problems. For example, homoclinic connections to q_ϵ (which has two-dimensional stable and unstable manifolds) as well as heteroclinic cycles involving p_ϵ and q_ϵ.

6. An application: modal dynamics of the damped, driven, nonlinear Schrödinger equation

In this section we use the general theory to study the dynamics of the two-mode model of the damped, driven nonlinear Schrödinger equation described in the introduction. We rewrite the two-mode equa-

tions given in the introduction below:

$$-\mathrm{i}\dot{c} + \left(\tfrac{1}{2}|c|^2 + \tfrac{1}{2}|b|^2 - 1\right)c + \tfrac{1}{2}(cb^* + c^*b)b = \mathrm{i}\epsilon\alpha c + \mathrm{i}\epsilon\Gamma,$$
$$-\mathrm{i}\dot{b} + \left[\tfrac{1}{2}|c|^2 + \tfrac{3}{4}|b|^2 - (1+k^2)\right]b + \tfrac{1}{2}(cb^* + c^*b)c = \mathrm{i}\epsilon\beta b. \qquad (6.1)_\epsilon$$

The unperturbed equations are obtained from $(6.1)_\epsilon$ by setting $\epsilon = 0$

$$-\mathrm{i}\dot{c} + \left(\tfrac{1}{2}|c|^2 + \tfrac{1}{2}|b|^2 - 1\right)c + \tfrac{1}{2}(cb^* + c^*b)b = 0,$$
$$-\mathrm{i}\dot{b} + \left(\tfrac{1}{2}|c|^2 + \tfrac{3}{4}|b|^2 - (1+k^2)\right)b + \tfrac{1}{2}(cb^* + c^*b)c = 0. \qquad (6.1)_0$$

It is straightforward to verify that the unperturbed equations possess the following two integrals:

$$H_0 = \tfrac{1}{8}|c|^4 + \tfrac{1}{2}|b|^2|c|^2 + \tfrac{3}{16}|b|^4 - \tfrac{1}{2}(1+k^2)|b|^2 - \tfrac{1}{2}|c|^2 + \tfrac{1}{8}(b^2c^{*2} + b^{*2}c^2) \qquad (6.2)$$

and

$$I = \tfrac{1}{2}(|c|^2 + |b|^2). \qquad (6.3)$$

the unperturbed system has the following Hamiltonian form:

$$\dot{c} = -2\mathrm{i}\frac{\partial H}{\partial c^*}, \quad \dot{b} = -2\mathrm{i}\frac{\partial H}{\partial b^*}. \qquad (6.4)$$

By inspection, we see that the unperturbed equations are invariant under the following coordinate transformations:

$$(c,b) \to (-c,b), \quad (c,b) \to (c,-b), \quad (c,b) \to (c,b)\,\mathrm{e}^{\mathrm{i}\chi}, \qquad (6.5\text{a, b, c})$$

where χ is any real number.

Eqs. (6.1) are not of the form $(1.1)_\epsilon$ where our theory can be applied. We can write them in the appropriate form by introducing the following coordinate transformation first used, to our knowledge, in [19]

$$c = |c|\,\mathrm{e}^{\mathrm{i}\gamma}, \quad b = (x + \mathrm{i}y)\,\mathrm{e}^{\mathrm{i}\gamma}. \qquad (6.6\text{a, b})$$

In these coordinates the unperturbed equations become

$$\dot{x} = -k^2y - \tfrac{3}{4}x^2y + \tfrac{1}{4}y^3 = \frac{\partial H_0}{\partial y},$$
$$\dot{y} = (k^2 - 2I)x + \tfrac{7}{4}x^3 + \tfrac{3}{4}xy^2 = -\frac{\partial H_0}{\partial x},$$
$$\dot{I} = 0 = \frac{\partial H_0}{\partial \gamma},$$
$$\dot{\gamma} = 1 - I - x^2 = -\frac{\partial H_0}{\partial I}. \qquad (6.7)_0$$

and the integrals are

$$H_0 = \tfrac{1}{2}I^2 - I - \tfrac{7}{16}x^4 - \tfrac{3}{8}x^2y^2 + \tfrac{1}{16}y^4 + \left(I - \tfrac{1}{2}k^2\right)x^2 - \tfrac{1}{2}k^2y^2 \quad \text{and} \quad I.$$

The perturbed equations are given by

$$\begin{aligned}
\dot{x} &= -k^2y - \tfrac{3}{4}x^2y + \tfrac{1}{4}y^3 + \epsilon\left[\Gamma\frac{y}{\sqrt{2I - x^2 - y^2}}\sin\gamma - \beta x\right] \\
\dot{y} &= (k^2 - 2I)x + \tfrac{7}{4}x^3 + \tfrac{3}{4}xy^2 - \epsilon\left[\Gamma\frac{x}{\sqrt{2I - x^2 - y^2}}\sin\gamma + \beta y\right], \\
\dot{I} &= -\epsilon\left[\Gamma\sqrt{2I - x^2 - y^2}\cos\gamma + (\beta - \alpha)(x^2 + y^2) + 2\alpha I\right], \\
\dot{\gamma} &= 1 - I - x^2 + \epsilon\Gamma\frac{1}{\sqrt{2I - x^2 - y^2}}\sin\gamma,
\end{aligned} \tag{6.7}_\epsilon$$

which have the general form

$$\begin{aligned}
\dot{x} &= \frac{\partial H_0}{\partial y} + \epsilon\frac{\partial H_1}{\partial y} - \epsilon\beta x, \quad \dot{y} = -\frac{\partial H_0}{\partial x} - \epsilon\frac{\partial H_1}{\partial x} - \epsilon\beta y \\
\dot{I} &= \epsilon\frac{\partial H_1}{\partial \gamma} - \epsilon 2\alpha I - \epsilon(\beta - \alpha)(x^2 + y^2), \quad \dot{\gamma} = -\frac{\partial H_0}{\partial I} - \epsilon\frac{\partial H_1}{\partial I},
\end{aligned} \tag{6.8}$$

where

$$H_1 = -\Gamma\sqrt{2I - x^2 - y^2}\sin\gamma. \tag{6.9}$$

Eqs. (6.7) are essentially of the form of $(1.1)_\epsilon$, however there is one (insignificant) difference. Namely, in (6.8) there is a minus sign in front of $\partial H_0/\partial I$ as opposed to the form given in $(1.1)_\epsilon$. Eq. (6.8) can be put in the form of $(1.1)_\epsilon$ by taking $-\gamma$ instead of γ as the angle conjugate to I or by interchanging x and y and taking the negative of the Hamiltonian. We use the form given in (6.8) so that we can directly compare our results with the results on the damped, driven sine-Gordon equation described in the introduction since this form does not effect the geometry of the invariant manifolds described in the general theory. We next describe the integrable structure of the unperturbed equations $(6.7)_0$.

6.1. *The unperturbed integrable structure*

We now show that $(6.7)_0$ has the invariant manifold structure described in the general theory. In order to do this we must consider the x–y component of $(6.7)_0$ which we rewrite below:

$$\dot{x} = -k^2y - \tfrac{3}{4}x^2y + \tfrac{1}{4}y^3, \quad \dot{y} = (k^2 - 2I)x + \tfrac{7}{4}x^3 + \tfrac{3}{4}xy^2. \tag{6.7}_{x,y}$$

Note that $(6.7)_{x,y}$ has a fixed point at $(x, y) = (0, 0)$ for all values of I, this is a result of the symmetry given by (6.5b). A simple linear stability analysis shows that $(x, y) = (0, 0)$ is a saddle point for $I > \frac{1}{2}k^2$. Moreover, an examination of the level set of the Hamiltonian that contains the origin, i.e., $H(x, y, I) -$

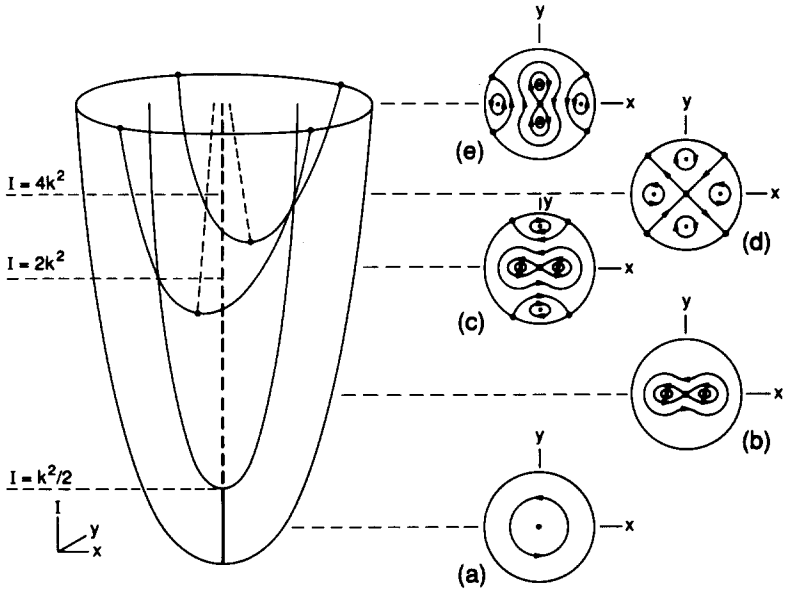

Fig. 6.1. Phase portraits of $(6.7)_{x,y}$ as a function of I. Note the global bifurcation at $I = 4k^2$.

$H(0,0,I) = 0$, shows that for each I in this range the origin has a pair of symmetric homoclinic orbits. Thus we have

$$\mathcal{M} = \{(x, y, I, \gamma) | x = y = 0,\ I > \tfrac{1}{2}k^2\} \tag{6.10}$$

and

$$W^s(\mathcal{M}) \cap W^s(\mathcal{M}) \equiv \{(x, y, I, \gamma) | H(x, y, I) - H(0,0,I) = 0\}. \tag{6.11}$$

The fact that one homoclinic orbit implies the existence a pair of homoclinic orbits follows from the symmetry (6.5b). In fig. 6.1 we show the phase portraits of $(6.7)_{x,y}$ as a function of I; note that at $I = 4k^2$ a global bifurcation associated with the homoclinic orbits connecting the origin occurs. In fig. 6.2 we show only the invariant manifold structure that will be important for our analysis.

Analytical expressions for the homoclinic orbits

In calculating the Melnikov functions it will be important to have analytical expressions for the homoclinic orbits of $(6.7)_{x,y}$ that connect the orign as a function of I. These calculations require a fair amount of tedious labor, however they are instructive so we include some of the details here.

We being by letting

$$x + iy = \sqrt{2B}\, e^{i\theta}. \tag{6.12}$$

In these coordinates the unperturbed equations $(6.7)_0$ become

$$\begin{aligned} &\dot{B} = -2B(I - B)\sin 2\theta, \quad \dot{\theta} = k^2 - I(1 + \cos 2\theta) + B(\tfrac{3}{2} + 2\cos 2\theta), \\ &\dot{I} = 0, \quad \dot{\gamma} = 1 - I - B(1 + \cos 2\theta), \end{aligned} \tag{6.13}$$

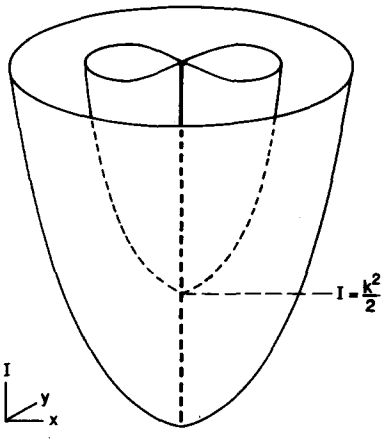

Fig. 6.2. The invariant manifold structure that plays a role in our analysis.

with the Hamiltonian H_0 being transformed into

$$\tilde{H}_0 = \tfrac{1}{2}I^2 - I - \tfrac{3}{4}B^2 + (I - k^2)B + B(I - B)\cos 2\theta. \tag{6.14}$$

By continuity, the value of the Hamiltonian on the orbits homoclinic to $\mathcal{M}$ is the same as the value of the Hamiltonian on $\mathcal{M}$, which is $\frac{1}{2}I^2 - I$. Equating this value to (6.14) and canceling the common factor B gives

$$B = I - \frac{4k^2 - I}{3 + 4\cos 2\theta}. \tag{6.15}$$

Substituting (6.15) into the $\dot{\theta}$ component of (6.13) gives the following equation for $\dot{\theta}$ on the homoclinic orbits:

$$\dot{\theta} = I(1 + \cos 2\theta) - k^2. \tag{6.16}$$

Next we let

$$\psi = \gamma + \theta \tag{6.17}$$

and add (6.16) and the $\dot{\gamma}$ component of (6.13) to get the following equation for $\dot{\psi}$ on the homoclinic orbits:

$$\dot{\psi} = 1 - I - \tfrac{1}{4}B. \tag{6.18}$$

Eqs. (6.15), (6.16), and (6.18) provide us with the necessary (and sufficiently simple) relationships for solving for the homoclinic orbits. Eq. (6.16) can be integrated directly to obtain $\theta(t)$ on the homoclinic orbits. This result can be substituted into (6.15) to obtain $B(t)$ on the homoclinic orbits, from which $x(t)$ and $y(t)$ can be obtained from (6.12). Eq. (6.18) can be integrated directly to give $\psi(t)$ on the homoclinic orbits. This result can then be added to $\theta(t)$ to give $\gamma(t)$ via (6.17). In computing these integrals there

will be two cases corresponding to $I > 4k^2$ and $\frac{1}{2}k^2 < I < 4k^2$. The fact that there are two cases stems from the global bifurcation at $I = 4k^2$ of the equation $(6.7)_{x,y}$ mentioned earlier.

$I > 4k^2$ with initial conditions $\theta(t=0) = \frac{1}{2}\pi$, $\psi(t=0) = \psi_0$. For this case we have

$$B = \frac{4k^2(2I - k^2)}{(I - 4k^2)\cosh\left(2k\sqrt{2I - k^2}\, t\right) + I + 3k^2}, \tag{6.19}$$

$$\cot\theta = \frac{k}{\sqrt{2I - k^2}} \tanh\left(k\sqrt{2I - k^2}\, t\right), \tag{6.20}$$

$$\psi = -\frac{1}{\sqrt{7}} \tanh^{-1}\left(\sqrt{\frac{7k^2}{2I - k^2}} \tanh\left(k\sqrt{2I - k^2}\, t\right)\right) + (1 - I)t + \psi_0. \tag{6.21}$$

$\frac{1}{2}k^2 < I < 4k^2$ with initial conditions $\theta(t=0) = 0$, $\psi(t=0) = \psi_0$. For this case we have

$$B = \frac{4k^2(2I - k^2)}{(4k^2 - i)\cosh\left(2k\sqrt{2I - k^2}\, t\right) + I + 3k^2}. \tag{6.22}$$

$$\tan\theta = \frac{\sqrt{2I - k^2}}{k} \tanh\left(k\sqrt{2I - k^2}\, t\right), \tag{6.23}$$

$$\psi = -\frac{1}{\sqrt{7}} \tanh^{-1}\left(\sqrt{\frac{2I - k^2}{7k^2}} \tanh\left(k\sqrt{2I - k^2}\, t\right)\right) + (1 - I)t + \psi_0. \tag{6.24}$$

Orbits homoclinic to the circle of fixed points

From $(6.7)_0$ and (6.10) we see that the unperturbed equations restricted to $\mathcal{M}$ are given by

$$\dot{I} = 0, \quad \dot{\gamma} = 1 - I. \tag{6.25}$$

hence, (6.25) has a resonance at $I = 1$. As discussed in section 2, this resonance is manifested as a circle of fixed points on $\mathcal{M}$ and we would like to use the expressions for the homoclinic orbits given above to compute the phase shift, $\Delta\gamma$, of orbits that are asymptotic to points on the circle of fixed points as $t \to \pm\infty$. There are two cases.

$0 < k < \frac{1}{2}$. For this case we obtain the following expressions from (6.20) and (6.21) with $I = 1$.

$$\psi(-\infty) = \frac{1}{\sqrt{7}} \tanh^{-1}\left(\sqrt{\frac{7k^2}{2 - k^2}}\right) + \psi_0, \quad \psi(\infty) = -\frac{1}{\sqrt{7}} \tanh^{-1}\left(\sqrt{\frac{7k^2}{2 - k^2}}\right) + \psi_0, \tag{6.26, 6.27}$$

$$\theta(-\infty) = \pi - \cot^{-1}\left(\frac{k}{\sqrt{2 - k^2}}\right), \quad \theta(\infty) = \cot^{-1}\left(\frac{k}{\sqrt{2 - k^2}}\right). \tag{6.28, 6.29}$$

Using (6.17) gives

$$\gamma(-\infty) = \gamma_0 - \tfrac{1}{2}\pi + \cot^{-1}\left(\frac{k}{\sqrt{2-k^2}}\right) + \frac{1}{\sqrt{7}}\tanh^{-1}\left(\sqrt{\frac{7k^2}{2-k^2}}\right), \tag{6.30}$$

$$\gamma(\infty) = \gamma_0 + \tfrac{1}{2}\pi - \cot^{-1}\left(\frac{k}{\sqrt{2-k^2}}\right) - \frac{1}{\sqrt{7}}\tanh^{-1}\left(\sqrt{\frac{7k^2}{2-k^2}}\right). \tag{6.31}$$

Hence, we have

$$\Delta\gamma \equiv \gamma(+\infty) - \gamma(-\infty) = \pi - 2\cot^{-1}\left(\frac{k}{\sqrt{2-k^2}}\right) - \frac{2}{\sqrt{7}}\tanh^{-1}\left(\sqrt{\frac{7k^2}{2-k^2}}\right). \tag{6.32}$$

In fig. 6.3 we plot $\Delta\gamma$ as a function of k for this case. From this plot we see that $\Delta\gamma$ is a monotonic function that approaches zero as k goes to zero and $-\infty$ as k goes to $\frac{1}{2}$.

$\frac{1}{2} < k < \sqrt{2}$. For this case, using (6.23) and (6.24) at $I = 1$ gives

$$\psi(-\infty) = \frac{1}{\sqrt{7}}\tanh^{-1}\left(\sqrt{\frac{2-k^2}{7k^2}}\right) + \psi_0, \quad \psi(\infty) = -\frac{1}{\sqrt{7}}\tanh^{-1}\left(\sqrt{\frac{2-k^2}{7k^2}}\right) + \psi_0, \tag{6.33, 6.34}$$

$$\theta(-\infty) = -\tan^{-1}\left(\frac{\sqrt{2-k^2}}{k}\right), \quad \theta(\infty) = \tan^{-1}\left(\frac{\sqrt{2-k^2}}{k}\right). \tag{6.35, 6.36}$$

Using (6.17) gives

$$\gamma(-\infty) = \gamma_0 + \tan^{-1}\left(\frac{\sqrt{2-k^2}}{k}\right) + \frac{1}{\sqrt{7}}\tanh^{-1}\left(\sqrt{\frac{2-k^2}{7k^2}}\right), \tag{6.37}$$

$$\gamma(\infty) = \gamma_0 - \tan^{-1}\left(\frac{\sqrt{2-k^2}}{k}\right) - \frac{1}{\sqrt{7}}\tanh^{-1}\left(\sqrt{\frac{2-k^2}{7k^2}}\right). \tag{6.38}$$

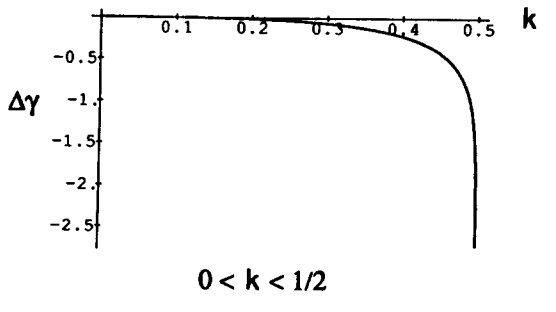

Fig. 6.3. The graph of $\Delta\gamma$ as a function of k for $0 < k < \frac{1}{2}$.

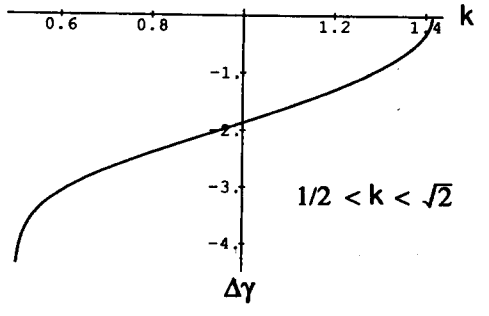

Fig. 6.4. The graph of $\Delta\gamma$ as a function of k for $\frac{1}{2} < k < \sqrt{2}$.

Hence we have

$$\Delta\gamma \equiv \gamma(+\infty) - \gamma(-\infty) = -2\tan^{-1}\left(\frac{\sqrt{2-k^2}}{k}\right) - \frac{2}{\sqrt{7}}\tanh^{-1}\left(\sqrt{\frac{2-k^2}{7k^2}}\right). \tag{6.39}$$

In fig. 6.4 we plot $\Delta\gamma$ as a function of k for this case. From this plot we see that $\Delta\gamma$ is a monotonic function that approaches $-\infty$ as k goes to $\frac{1}{2}$ and zero as k goes to $\sqrt{2}$. Note that the function is always negative.

6.2. *Dynamics near the resonance on* $\mathscr{A}_\epsilon$

We now want to consider the dynamics of the perturbed vector field restricted to $\mathscr{A}_\epsilon$ near the resonance at $I = 1$. For this example, the perturbed normally hyperbolic invariant manifold coincides with the unperturbed manifold since $x = y = 0$ is invariant under the perturbed dynamics also (this is related to the fact that the symmetry (6.5b) is also present for the perturbed problem); hence, the general theory described in section 3 simplifies considerably. Thus the perturbed vector field $(6.7)_\epsilon$ restricted to $\mathscr{A}_\epsilon$ is given by

$$\dot{I} = -\epsilon\left(\Gamma\sqrt{2I}\cos\gamma + 2\alpha I\right), \quad \dot{\gamma} = 1 - I + \frac{\epsilon\Gamma}{\sqrt{2I}}\sin\gamma. \tag{6.40}$$

We will be interested in the dynamics of (6.40) near $I = 1$, i.e., in an annulus on the cylinder with coordinates I–γ centered at $I = 1$. Following section 3.2, we introduce the coordinate transformation $I = 1 + \sqrt{\epsilon}\,h$, rescale time by letting $\tau = \sqrt{\epsilon}\,t$, and Taylor expand in $\sqrt{\epsilon}$ about $\sqrt{\epsilon} = 0$. Eqs. (6.40) then become

$$h' = -\sqrt{2}\,\Gamma\cos\gamma - 2\alpha - \sqrt{\epsilon}\left(2\alpha + \frac{\Gamma}{\sqrt{2}}\cos\gamma\right)h + \mathscr{O}(\epsilon), \quad \gamma' = -h + \sqrt{\epsilon}\,\frac{\Gamma}{\sqrt{2}}\sin\gamma + \mathscr{O}(\epsilon). \tag{6.41}$$

where the prime denotes the differentiation with respect to τ.

In the limit $\epsilon \to 0$ (6.41) becomes

$$h' = -\Gamma\sqrt{2}\cos\gamma - 2\alpha, \quad \gamma' = -h. \tag{6.42}$$

Eq. (6.42) is just the familiar equation for a pendulum subject to a constant torque. Moreover, it is Hamiltonian with Hamiltonian function given by

$$\mathscr{H} = \tfrac{1}{2}h^2 - \Gamma\sqrt{2}\sin\gamma - 2\alpha\gamma. \tag{6.43}$$

A simple analysis shows that (6.42) has two fixed points whose coordinates must satisfy

$$h = 0, \quad \cos\gamma = -\sqrt{2}\,\frac{\alpha}{\Gamma}. \tag{6.44}$$

Hence, the (h, γ) coordinates of the two fixed points are given by

$$p_0 = \left(0, \pi - \cos^{-1}\sqrt{2}\,\frac{\alpha}{\Gamma}\right), \quad q_0 = \left(0, \pi + \cos^{-1}\sqrt{2}\,\frac{\alpha}{\Gamma}\right) \tag{6.45}$$

where p_0 is a center and q_0 is a saddle. These two fixed points coalesce in a saddle-node bifurcation at $\alpha = \Gamma/\sqrt{2}$. Moreover, the saddle point is connected to itself by a homoclinic orbit. In fig. 6.5 we show the phase portrait of (6.42). Note that the length (measured with respect to the γ coordinate) of the "fish" in fig. 6.5 varies monotonically with α (from 0 to 2π) and the maximum width (measured with respect to the h coordinate) varies linearly with Γ.

Next we want to determine what becomes of the phase portrait when the higher order terms in $\sqrt{\epsilon}$ are taken into account. The trace of the linearization of (6.41) is given by $-2\sqrt{\epsilon}\,\alpha + \mathcal{O}(\epsilon)$. Hence, by lemma 3.2, for ϵ sufficiently small and $\alpha > 0$, p_0 becomes a sink and the homoclinic orbit breaks with a branch of the unstable manifold of q_ϵ falling into p_ϵ as shown in fig. 6.6.

6.3. *Calculation of the Melnikov function*

Using (4.1), the integrand of the Melnikov function is given by

$$\begin{aligned}
&\frac{\partial H_0}{\partial x}\frac{\partial H_1}{\partial y} - \frac{\partial H_0}{\partial y}\frac{\partial H_1}{\partial x} \\
&\quad + \left(\frac{\partial H_0}{\partial I}(x, y, I) - \frac{\partial H_0}{\partial I}(0, 0, I)\right)\frac{\partial H_1}{\partial \gamma} \\
&\quad - \beta\left(x\frac{\partial H_0}{\partial x} + y\frac{\partial H_0}{\partial y}\right) \\
&\quad - \left(\frac{\partial H_0}{\partial I}(x, y, I) - \frac{\partial H_0}{\partial I}(0, 0, I)\right) \\
&\quad \times \left[2\alpha I + (\beta - \alpha)(x^2 + y^2)\right],
\end{aligned} \tag{6.46}$$

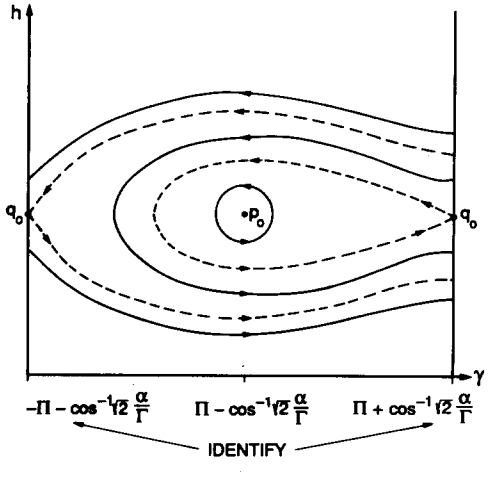

Fig. 6.5. The phase portrait of eq. (6.42) (for a typical value of Γ and α).

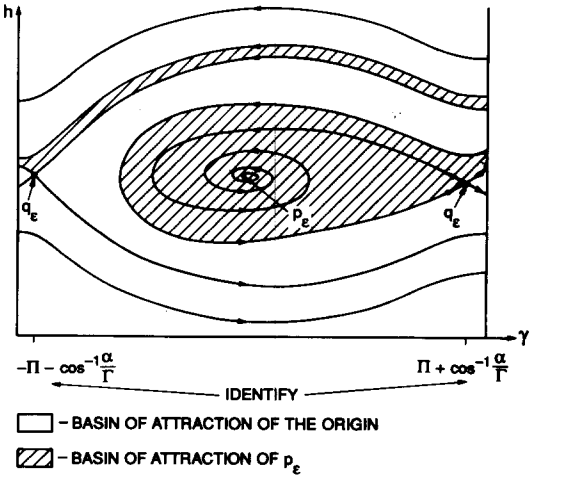

Fig. 6.6. The phase portrait of eq. (6.41) (for a typical value of Γ and α).

where the first line is due to the constant driving and the second line is due to the dissipation. Recall that the integrand of the Melnikov function will be integrated along the unperturbed heteroclinic orbit that is asymptotic to p_0 as $t \to -\infty$ which is obtained from the general expressions given in section 6.1 by setting $I = 1$ and $\gamma_0 = \pi - \cos^{-1}\sqrt{2}(\alpha/\Gamma)$. In the formulae that we derive we will explicitly denote the arguments of the $\partial H_0/\partial I$ term since this term may be evaluated on either $\mathcal{M}$ or on an orbit homoclinic to $\mathcal{M}$. The Melnikov function integrand can be simplified considerably. A simple calculation using the chain rule gives

$$\frac{\partial H_0}{\partial x}\frac{\partial H_1}{\partial y} - \frac{\partial H_0}{\partial y}\frac{\partial H_1}{\partial x} = -\frac{dH_1}{dt} - \frac{\partial H_1}{\partial \gamma}\frac{\partial H_0}{\partial I}, \tag{6.47}$$

where we have used the fact that $\dot{I} = 0$. The Melnikov function thus simplifies to

$$\begin{aligned}&\frac{-\mathrm{d}H_1}{\mathrm{d}t} - \frac{\partial H_0}{\partial I}(0,0,I)\frac{\partial H_1}{\partial \gamma} - \beta(\dot{x}y - \dot{y}x)\\ &\quad - \left(\frac{\partial H_0}{\partial I}(x,y,I) - \frac{\partial H_0}{\partial I}(0,0,I)\right)\left[2\alpha I + (\beta-\alpha)(x^2+y^2)\right],\end{aligned} \tag{6.48}$$

where we have used

$$\dot{x} = \frac{\partial H_0}{\partial y}, \quad \dot{y} = -\frac{\partial H_0}{\partial x}.$$

Up to this point, the expression given in (6.48) is completely general; we have not set $I = 1$ or fixed γ_0 at the value corresponding to p_0. Now we set $I = 1$ so that $(\partial H_0/\partial I)(0,0,I) = 0$ in which case (6.48) becomes

$$\frac{-\mathrm{d}H_1}{\mathrm{d}t} + \beta(x\dot{y} - y\dot{x}) + 2\alpha\dot{\gamma} + (\beta-\alpha)(x^2+y^2)\dot{\gamma}, \tag{6.49}$$

where we have used

$$\dot{\gamma} = -\frac{\partial H_0}{\partial I}.$$

We now integrate (6.49) around the unperturbed heteroclinic orbit at $I = 1$ that approaches p_0 asymptotically as $t \to -\infty$. We examine each term in (6.49) individually.

It is clear that the first term in (6.49) can be integrated directly to give

$$-\int_{-\infty}^{+\infty}\frac{\mathrm{d}H_1}{\mathrm{d}t}\,\mathrm{d}t = \sqrt{2}\Gamma[\sin\gamma(+\infty) - \sin\gamma(-\infty)]. \tag{6.50}$$

Recalling from section 2.4 that $\Delta\gamma \equiv \gamma(+\infty) - \gamma(-\infty)$ and using trigonometric identities allows us to simplify (6.50) to

$$-\int_{-\infty}^{+\infty}\frac{\mathrm{d}H_1}{\mathrm{d}t}\,\mathrm{d}t = \sqrt{2}\Gamma[\sin\gamma(-\infty)(\cos\Delta\gamma - 1) + \cos\gamma(-\infty)\sin\Delta\gamma]. \tag{6.51}$$

From (6.44) and (6.45) we have

$$\sin\gamma(-\infty) = \sqrt{1-2\left(\frac{\alpha}{\Gamma}\right)^2}, \quad \cos\gamma(-\infty) = -\sqrt{2}\,\frac{\alpha}{\Gamma}, \tag{6.52}$$

which substituted into (6.51) gives

$$-\int_{-\infty}^{+\infty}\frac{dH_1}{dt}\,dt = \sqrt{2}\Gamma\left[\sqrt{1-2\left(\frac{\alpha}{\Gamma}\right)^2}(\cos\Delta\gamma - 1) - \sqrt{2}\,\frac{\alpha}{\Gamma}\sin\Delta\gamma\right]. \tag{6.53}$$

It is also easy to see that the third term in (6.49) can be integrated directly to give

$$2\alpha\int_{-\infty}^{+\infty}\dot{\gamma}\,dt = 2a\Delta\gamma. \tag{6.54}$$

We now examine the second and fourth terms in (6.49). Differentiating (6.12) with respect to time, one can easily show that

$$x\dot{y} - y\dot{x} = 2B\dot{\theta} \tag{6.55}$$

and, taking the modulus of (6.12) gives

$$x^2 + y^2 = 2B. \tag{6.56}$$

From these two relations and (6.17) we obtain

$$\dot{\theta} = \dot{\psi} - \dot{\gamma} = \frac{x\dot{y} - y\dot{x}}{x^2+y^2} \tag{6.57}$$

or

$$\dot{\gamma} = \dot{\psi} - \frac{x\dot{y} - y\dot{x}}{x^2+y^2}. \tag{6.58}$$

Substituting (6.58) into the fourth term of (6.49) and combining the result with the second term of (6.49) gives

$$\beta(x\dot{y} - y\dot{x}) + (\beta-\alpha)(x^2+y^2)\dot{\gamma} = \alpha(x\dot{y} - y\dot{x}) + (\beta-\alpha)(x^2+y^2)\dot{\psi}. \tag{6.59}$$

We next integrate the two terms in (6.59) around the heteroclinic orbit. Using (6.55), the integral of the first term of (6.59) becomes

$$\int_{-\infty}^{+\infty}(x\dot{y} - y\dot{x})\,dt = 2\int_{-\infty}^{+\infty}B\dot{\theta}\,dt. \tag{6.60}$$

Using (6.15), at $I = 1$ we have

$$B = 1 - (4k^2 - 1)\frac{1}{3 + 4\cos\theta}$$

which, when substituted into (6.60), gives

$$\int_{-\infty}^{+\infty} (x\dot{y} - y\dot{x})\,\mathrm{d}t = 2\Delta\theta - 2(4k^2 - 1)\int_{\theta(-\infty)}^{\theta(+\infty)} \frac{\mathrm{d}\theta}{3 + 4\cos\theta}, \tag{6.61}$$

where

$$\Delta\theta \equiv \theta(+\infty) - \theta(-\infty)$$

can be obtained from (6.28) and (6.29) for $0 < k < \frac{1}{2}$ and (6.35) and (6.36) for $\frac{1}{2} < k < \sqrt{2}$. Integrating the last term in (6.61) gives

$$\int_{-\infty}^{+\infty} (x\dot{y} - y\dot{x})\,\mathrm{d}t = 2\Delta\theta - 2(4k^2 - 1)\Delta\psi, \tag{6.62}$$

where $\Delta\psi \equiv \psi(+\infty) - \psi(-\infty)$ and from (6.26), (6.27), and (6.34) we have

$$\Delta\psi = -\frac{2}{\sqrt{7}} \tanh^{-1}\sqrt{\frac{7k^2}{2 - k^2}} \quad \text{for } 0 < k < \tfrac{1}{2}, \quad \text{or}$$

$$\Delta\psi = -\frac{2}{\sqrt{7}} \tanh^{-1}\sqrt{\frac{2 - k^2}{7k^2}} \quad \text{for } \tfrac{1}{2} < k < \sqrt{2}. \tag{6.63}$$

Finally, we integrate the last term in (6.59) by first using (6.56) and (6.18) to obtain

$$\int_{-\infty}^{+\infty} (x^2 + y^2)\dot{\psi}\,\mathrm{d}t = -\frac{1}{2}\int_{-\infty}^{+\infty} B^2\,\mathrm{d}t. \tag{6.64}$$

Using the expressions for $B(t)$ in (6.19) for $0 < k < \frac{1}{2}$ and (6.22) for $\frac{1}{2} < k < \sqrt{2}$ gives

$$-\tfrac{1}{2}\int_{-\infty}^{+\infty} B^2\,\mathrm{d}t = \tfrac{8}{7}\left[k\sqrt{2 - k^2} + (1 + 3k^2)\,\Delta\psi\right], \tag{6.65}$$

where $\Delta\psi$ is given in (6.63).

Using (6.53), (6.54), (6.59), (6.62), (6.64), and (6.65), the Melnikov function becomes

$$M(a, b, k) = \left[\sqrt{1 - a^2}\,(\cos\Delta\gamma - 1) - a\sin\Delta\gamma\right] + 4ak^2\,\Delta\psi + (b - a)\tfrac{4}{7}\left[k\sqrt{2 - k^2} + (1 + 3k^2)\Delta\psi\right], \tag{6.66}$$

where

$$a = \sqrt{2}\,\frac{\alpha}{\Gamma}, \quad b = \sqrt{2}\,\frac{\beta}{\Gamma}. \tag{6.67}$$

6.4. The existence of an orbit homoclinic to p_ϵ

Following the theory developed in section 4, in order to show that there exists an orbit homoclinic to p_ϵ we must first show that the Melnikov function has a simple zero. This condition is a sufficient condition for the existence of an orbit that is asymptotic to p_ϵ as $t \to -\infty$ and asymptotic to an orbit in $\mathscr{A}_\epsilon$ as $t \to +\infty$. In order to verify that this orbit is asymptotic to an orbit in $\mathscr{A}_\epsilon$ that approaches p_ϵ as $t \to +\infty$ it is sufficient to show that the unperturbed heteroclinic orbit that is asymptotic to p_0 as $t \to -\infty$ returns to the circle of fixed points as $t \to +\infty$ at a γ value that places it within the unperturbed homoclinic orbit in $\mathscr{A}_\epsilon$ (in the rescaled coordinates) that connects q_0. In order to simplify the calculations we take $k = 1$. This is also the value of k taken in the numerical experiments on the damped, driven sine–Gordon equation that was the motivation for this work described in the introduction. *Also, we note that in the following we round all numerical quantities (i.e., $\Delta\gamma$ and $\Delta\psi$ at $k = 1$) to two decimal places. In practice, these can be computed to any desired degree of accuracy, however, recall that the relevant quantities are only $\mathscr{O}(\epsilon)$ or $\mathscr{O}(\sqrt{\epsilon})$ approximations to the actual quantities for the perturbed system.*

Setting $k = 1$, and using (6.39), (6.63), and (6.66), the condition that $M(a, b, 1) = 0$ is given by

$$b = -11.20\sqrt{1 - a^2} - 1.14a. \tag{6.68}$$

Thus it follows that for all values of $a \in (0, 1)$ there is a value of b (which is always negative) such that the Melnikov function is zero. Moreover, it is an easy calculation to show that this zero is a simple zero.

Next, we show that the unstable manifold of p_ϵ returns to $\mathscr{A}_\epsilon$ at the appropriate location. From (6.39), at $k = 1$, we have

$$\Delta\gamma = -1.87. \tag{6.69}$$

The location of the noise of the fish is given by the solution of the following transcendental equation (cf. eq. (3.13)):

$$\mathscr{H}(0, \gamma_n) - \mathscr{H}(0, \pi + \cos^{-1} a) = \sin(\pi + \cos^{-1} a) - \sin\gamma_n + a\big(\pi + \cos^{-1} a - \gamma_n\big) = 0, \tag{6.70}$$

where $\mathscr{H}$ is given by (6.43) and $a = \sqrt{2}\,\alpha/\Gamma$. From our knowledge of the phase portrait of the pendulum with torque, we know that (6.70) has exactly two solutions, one being the γ value corresponding to q_0, denoted γ_{q_0}. This equation can easily be solved numerically using Newton's method. In fig. 6.7 we present γ_n, γ_{q_0}, and $\gamma_{p_0} + \Delta\gamma$ for values of a between 0 and 1 (note: γ_{p_0} is the γ coordinate of p_0).

Recall that $\Delta\gamma$ only depends on k; in particular, it does not depend on a or b. Also, γ_{q_0} and γ_{p_0} depend only on a. These facts allow for the simple graphical depiction in fig. 6.7 of the criteria for the unstable manifold of p_ϵ to asymptote to an orbit inside the "fish" on $\mathscr{A}_\epsilon$ as given in theorem 4.2. Thus, we see that for $a \in (0, 0.19)$ we have $\gamma_n < \gamma_{p_0} + \Delta\gamma < \gamma_{q_0}$. Hence, we have the following theorem.

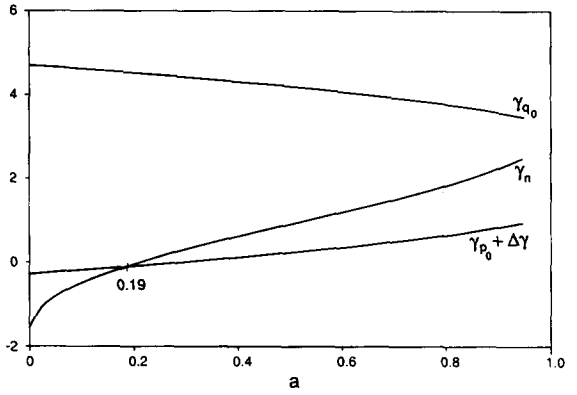

Fig. 6.7. Graphs of γ_{q_0}, γ_n, and $\gamma_{p_0}+\Delta\gamma$ as a function of $a=\sqrt{2}\alpha/\Gamma$.

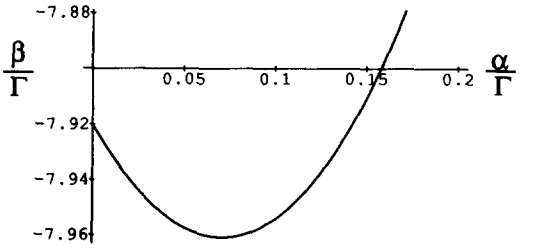

Fig. 6.8. Parameter values in the $\beta/\Gamma-\alpha/\Gamma$ plane (with $k=1$) for which orbits homoclinic to p_ϵ exist, $\alpha/\Gamma<0.13$.

Theorem 6.1. For ϵ sufficiently small, $\alpha/\Gamma<0.13$, and $\beta/\Gamma\approx -7.92\sqrt{1-2(\alpha/\Gamma)^2}-1.14\alpha/\Gamma$, p_ϵ has a homoclinic orbit.

The "approximate" sign in theorem 6.1 refers to the fact that the Melnikov function is only an accurate measure of distance up to $\mathcal{O}(\epsilon^2)$ and the numerical quantities have only been given to two decimal places. In fig. 6.8 we show the curve near which the homoclinic orbit occurs. It is a simple matter to show that the hypotheses of theorem 5.1 are satisfied for this system (see [15]). In particular, from remark 2 following theorem 5.1, we need only show that the two $\mathcal{O}(1)$ (real) eigenvalues associated with the linearization of the perturbed vector field about p_ϵ are not equal in magnitude. This is an easy calculation since the 4×4 Jacobian matrix is block diagonal (with two 2×2 blocks). At $k=1$ the two $\mathcal{O}(1)$ eigenvalues are found to be $\pm 1-\epsilon\beta+\mathcal{O}(\epsilon^2)$. Hence chaos, in the sense of theorem 5.1, occurs for parameter values near this curve. We describe this more fully in the next section.

6.5. *The geometrical interpretation of chaos in phase space*

Theorem 5.1 implies that the orbit structure near the homoclinic orbits is chaotic in the sense that a three-dimensional Poincaré map defined near the homoclinic orbits has an invariant Cantor set on which the dynamics is topologically conjugate to a shift map acting on a countable set of symbols. As stated, this result is not sufficient to explain how this chaos is manifested in terms of the motion in phase space of the dynamical system. For this, one must go into a bit of detail on the construction of the Poincaré map near the homoclinic orbits, from this one can explain geometrically the meaning of the chaotic motion implied by the symbolic dynamics. In this section we will outline the geometrical construction of the Poincaré map and the associated Smale horseshoe and symbolic dynamics.

We begin by expressing the vector field in coordinates localized about p_ϵ as follows:

$$\dot{x}=-\rho x-\omega y+F^x(x,y,z,w),\quad \dot{y}=\omega x-\rho y+F^y(x,y,z,w),$$
$$\dot{z}=-\lambda z+F^z(x,y,z,w),\quad \dot{w}=\nu w+F^w(x,y,z,w), \tag{6.71}$$

where we are assuming that the parameters and ϵ are fixed (and, hence, not explicitly shown) at the appropriate values for the existence of a homoclinic orbit to p_ϵ. The coordinate transformations taking

the vector field into the form (6.71) involve translating the fixed point to the origin followed by a linear transformation that puts the linear part in real Jordan canonical form; such results are standard and can be found in, e.g., [7]. The coordinates (x, y) can be thought of as coordinates on $\mathscr{A}_\epsilon$ and (6.71) is unchanged under the coordinate transformation $(x, y, z, w) \to (x, y, -z, -w)$ which is the manifestation of the symmetry (6.5b) in these coordinates. Hence, if the origin has one homoclinic orbit then it must have another, which is just its image under this coordinate transformation.

The method of analysis of the orbit structure near homoclinic orbits is standard. Cross-sections transverse to the vector field are defined by

$$\begin{aligned}
\Pi_0^+ &= \{(x, y, z, w) | \delta e^{-2\pi\rho/\omega} \le x \le \delta,\ y = 0,\ 0 < w \le \delta,\ -\delta \le z \le \delta\},\\
\Pi_0^- &= \{(x, y, z, w) | \delta e^{-2\pi\rho/\omega} \le x \le \delta,\ y = 0,\ -\delta \le w < 0,\ -\delta \le z \le \delta\},\\
\Pi_1^+ &= \{(x, y, z, w) | w = \delta\} \quad \Pi_1^- = \{(x, y, z, w) | w = -\delta\}.
\end{aligned} \tag{6.72}$$

and are shown in fig. 6.9. As a convenient notation we define $\Pi_0 \equiv \Pi_0^+ \cup \Pi_0^-$. The Poincaré map will map Π_0 into Π_0 and will be constructed as the composition of two maps; a map of Π_0^+ into Π_1^+, denoted P_0^+ (resp. Π_0^- into Π_1^-, denoted P_0^-) which, for δ sufficiently small, is essentially given by the flow generated by the linearized vector field and a map of Π_1^+ into Π_0^+, denoted P_1^+ (resp. Π_1^- into Π_0^-, denoted P_1^-) which, for δ sufficiently small, is essentially an affine transformation around the homoclinic orbits outside a neighborhood of the fixed point. The Poincaré map P: $\Pi_0 \to \Pi_0$ is thus defined as $P \equiv P_1^+ \circ P_0^+ \cup P_1^- \circ P_0^-$, see fig. 6.10.

Now we want to show how a Smale horseshoe is constructed for this map. Consider the regions defined by

$$\begin{aligned}
R_k^+ &= \{(x, y, z, w) | \delta e^{-2\pi\rho/\omega} \le x \le \delta,\ -\delta < z \le \delta,\ \delta e^{-2\pi(k+1)\nu/\omega} \le w \le \delta e^{-2\pi k\nu/\omega}\},\\
R_k^- &= \{(x, y, z, w) | \delta e^{-2\pi\rho/\omega} \le x \le \delta,\ -\delta < z \le \delta,\ -\delta e^{-2\pi(k+1)\nu/\omega} \ge w \ge -\delta e^{-2\pi k\nu/\omega}\}.
\end{aligned} \tag{6.73}$$

and shown in fig. 6.11 for k fixed and sufficiently large. Under the map P, the image of R_k^+ intersects R_k^+ as shown in fig. 6.12, similarly for the image of R_k^- under P (just use the symmetry). Most importantly, $P(R_k^+)$ intersects both R_k^+ and R_k^-, similarly for $P(R_k^-)$. Hence, we can find smaller regions in R^+ and R^-, denoted H^+ and H^-, respectively, that are stretched, contracted, and mapped over themselves as shown in fig. 6.13. The invariant Cantor set is the intersection of all the forward and

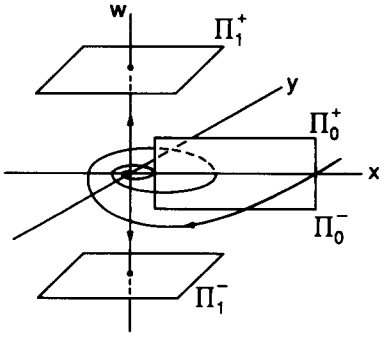

Fig. 6.9. The cross sections Π_0^+, Π_0^-, Π_1^+, and Π_1^- (with the z coordinate suppressed).

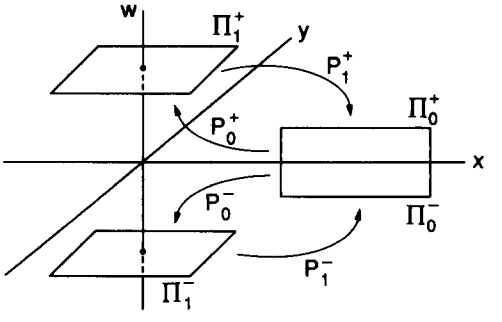

Fig. 6.10. The maps P_0^+, P_0^-, P_1^+, and P_1^- (with the z coordinate suppressed).

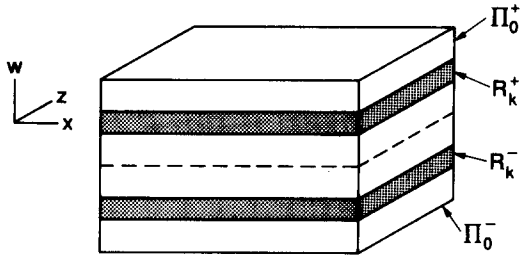

Fig. 6.11. The regions R_k^+ and R_k^- (with the y coordinate suppressed).

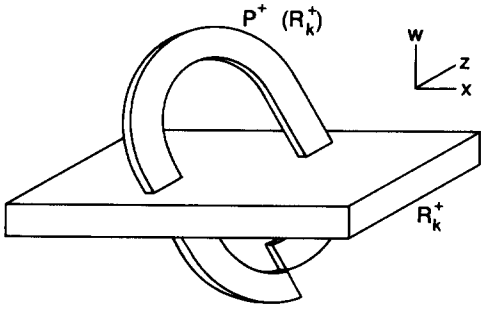

Fig. 6.12. The geometry of the intersection of R_k^+ and its image $P^+(R_k^+)$.

backward iterates of H^+ and H^- and the symbols in the symbolic dynamics can be taken to be + and −. Thus, in terms of motion in the phase space, a + in a symbol sequence corresponds to motion close to the homoclinic orbit having positive z and w coordinates and a − in a symbol sequence corresponds to motion close to the homoclinic orbit having negative z and w coordinates. Therefore, the chaos is manifested as random "jumping" around the two homoclinic orbits.

6.6. *On the relationship to the numerical experiments of Bishop et al.*

For the two-mode truncation of the damped and driven nonlinear Schrödinger equation studied in this section we have found a mechanism for chaotic dynamics that is similar to that observed by Bishop et al. Namely, the Silnikov mechanism gives rise to deterministic "random" jumping between two different types of spatially dependent states with an intermediate passage through a spatially independent, or flat state (the c-plane). However, the conditions under which this occurs are different than that studied by Bishop et al. in that they have equal damping in all of the modes (i.e., $\alpha = \beta$). Our methods show that no perturbative, homoclinic orbits to p_ϵ, of the type described in this paper, exist for this case. Moreover, in order for such a homoclinic orbit to exist the spatially dependent mode (the b variable) must be negatively damped (which is unphysical). The meaning of these results (if any) for the dynamics of the damped and driven nonlinear Schrödinger equation is unclear.

The original purpose for studying the two-mode truncation was as a "geometrically correct" model for the phase space structure of the nonlinear Schrodinger equation. With this point of view in mind, there

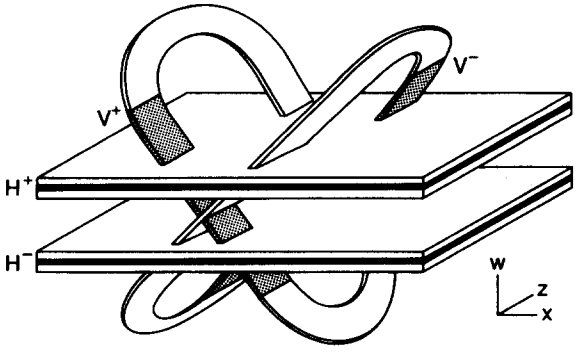

Fig. 6.13. The three-dimensional horseshoe map.

are additional mechanisms in the two-mode truncation that may give rise to the type of chaos observed in the numerical experiments of Bishop et al.; in particular, orbits homoclinic to q_ϵ. Preliminary work in collaboration with David McLaughlin, Ed Overman, and C. Xiong indicates that such orbits exist when both modes are negatively damped. This is work in progress that will be reported elsewhere. In any case, a theoretical description of the numerical experiments of Bishop et al. is still an open question.

Acknowledgements

The authors would like to express their gratitude to Nick Ercolani, Greg Forest, Dana Hobson, Phil Holmes, Tasso Kaper, and Pat Sethna for useful conversations. Zaichun Feng also made some useful comments and showed us some shortcuts and tricks for evaluating the Melnikov integrals. Dave McLaughlin has played an important role throughout the development of this theory. Discussions with him have greatly clarified our results. Ed Overman and C. Xiong's numerical simulations have also provided much insight and direction.

This research has been supported by an NSF Presidential Young Investigator Award and an ONR Young Investigator Award.

References

[1] J.C. Alexander and D.-Y. Cai, J. Math. Biol. 29 (1991) 405.
[2] N. Aubry, P. Holmes, J.L. Lumley and E. Stone, J. Fluid Mech. (1988) 115.
[3] T. Matsumoto, L.O. Chua and M. Komuro, IEEE Trans. Circ. Sys. 32 (1985) 798.
[4] D. David, D.D. Holm, and M.V. Tratnik, Phys. Lett. A 137 (1989) 355.
D. David, D.D. Holm, and M.V. Tratnik, Phys. Lett. A 138 (1989) 29.
[5] F.C. Moon, Chaotic Vibrations: An Introduction for Applied Scientists and Engineers (Wiley, New York, 1987).
[6] P. Gray and S.K. Scott, Chemical Oscillations and Instabilities: Non-Linear Chemical Kinetics (Oxford Univ. Press, Oxford, 1990).
[7] S. Wiggins, Global Bifurcations and Chaos – Analytical Methods (Springer, Berlin, 1988).
[8] A.R. Bishop, M.G. Forest, D.W. McLaughlin and E.A. Overman II, Physica D 23 (1986) 293;
E.A. Overman II, D.W. McLaughlin and A.R. Bishop, Physica D 19 (1986) 1;
A.R. Bishop, D.W. McLaughlin, M.G. Forest and E.A. Overman II, Phys. Lett. A 127 (1988) 335.
[9] A.R. Bishop, M.G. Forest, D.W. McLaughlin and E.A. Overman II, Phys. Lett. A 144 (1990) 17;
A.R. Bishop, R. Flesch, M.G. Forest, D.W. McLaughlin and E.A. Overman II, SIAM J. Math. Anal. 21 (1990) 1511.
[10] N. Ercolani, M.G. Forest and D.W. McLaughlin, Physica D 43 (1990) 349;
N. Ercolani, M.G. Forest and D.W. McLaughlin, Geometry of the modulational instability part II: global results, to appear;
N. Ercolani, M.G. Forest and D.W. McLaughlin, Geometry of the modulational instability. Part I: Local analysis, to appear.
[11] A.C. Newell, Solitons in Mathematics and Physics (SIAM, Philadelphia, 1985).
[12] N. Fenichel, Ind. Univ. Math. J. 21 (1971) 193;
N. Fenichel, Ind. Univ. Math. J. 23 (1974) 1109;
N. Fenichel, Ind. Univ. Math. J. 26 (1977) 81.
[13] S. Wiggins, Introduction to Applied Nonlinear Dynamical Systems and Chaos (Springer, Berlin, 1990).
[14] A.D. Morozov and L.P. Silnikov, Prikl. Mat. Mekh. USSR 47 (1984) 327.
[15] G. Kovačič, Chaos in a model of the forced and damped sine-Gordon equation, Caltech Ph.D. thesis.
[16] B.D. Greenspan and P.J. Holmes, SIAM J. Math. Anal. 15 (1984) 69.
[17] N. Fenichel, J. Diff. Eq. 31 (1979) 53.
[18] L.P. Silnikov, Math. USSR Sb. 10 (1970) 91.
[19] J.A. Zufiria, J. Fluid Mech. 191 (1988) 341.